

Nonlinear diffusion filtering of images using the topological gradient approach to edges detection

Monika Muszkiet

Vom Fachbereich Mathematik
der Technische Universität Kaiserslautern
zur Verleihung des akademischen Grades
Doktor der Naturwissenschaften
(Doctor rerum Naturalium, Dr. rer. nat.)
genehmigte Dissertation

1. Gutachter: Prof. Dr. Helmut Neunzert
2. Gutachter: Prof. Dr. Mohamed Masmoudi

Vollzug der Promotion: Kaiserslautern, 30 August 2007

Acknowledgment

First of all, I would like to thank Prof. Helmut Neunzert for the help and strong support during my PhD study. I am very grateful to him for all the time that he devoted to discussions and for turning my interest into problems of image processing. Without his understanding and encouragement, this work would not be finished.

Secondly, I want to thank PD Dr. Oleg Iliev for the help on numerics and for many practical advices. I am thankful to him for arranging my meeting with Prof. Raytcho Lazarov and Prof. Achi Brandt.

Next, I want to thank Prof. Mohamed Masmoudi for giving me idea of application the topological asymptotic analysis to problem of image processing, and Prof. Otmar Scherzer for useful advices and opportunity of working for one month together with his research group at University in Innsbruck.

Particular thanks to Prof. Wojciech Okrasinski for encouragement and introducing me to the field of industrial mathematics.

I would also like to thank the Fraunhofer Society and partially the Image Processing Department of the Fraunhofer Institute for Industrial Mathematics in Kaiserslautern for providing the financial support.

Finally, I would like to thank my family and friends for helping me to get free from thinking about work. A special thanks to Alfonso Caiazzo, Anna Naumovich, Iryna Rybak and Zahra Lakdawala for lot of adventures which made my stay in Kaiserslautern more colorful. Moreover, I am grateful to Zahra Lakdawala for correcting my english language in this thesis.

Contents

1	Introduction	1
2	An overview of diffusion filters	7
2.1	Linear diffusion filter	7
2.2	Nonlinear isotropic diffusion filter	8
2.2.1	The Perona and Malik model	8
2.2.2	Regularization of the Perona and Malik model	12
2.3	Nonlinear anisotropic diffusion filter	14
2.3.1	The Weickert model	14
3	Topological sensitivity analysis	17
3.1	Formulation of the problem	18
3.2	Asymptotic analysis	23
3.2.1	Preliminaries	23
3.2.2	Integral representation formula	25
3.2.3	Estimation of the Green function	28
3.2.4	Asymptotic expansion of $u_\varepsilon - u_0$ on the boundary of the inhomogeneity	34
3.2.5	Topological asymptotic expansion	41
3.3	Explicit formula of the topological gradient	46
3.4	Remarks	49
3.4.1	The generalized adjoint method	50
3.4.2	The domain truncation method	52

3.4.3	Formula of the ‘topological gradient’ derived using the ‘generalized adjoint method’	54
4	Finite volume discretization	57
4.1	Grid coarsening technique	58
4.2	Discretization of the Catté et al. model	59
4.2.1	Time discretization	60
4.2.2	Space discretization on an uniform grid	60
4.2.3	Space discretization on a nonuniform grid	62
4.3	Discretization of the Weickert model	65
4.3.1	Time discretization	65
4.3.2	Space discretization on an uniform grid	66
4.3.3	Space discretization on a nonuniform grid	68
4.4	Discretization of the first and the second order derivatives . .	72
4.4.1	Discretization on an uniform grid	72
4.4.2	Discretization on a nonuniform grid	73
5	Numerical experiments	75
5.1	Edge detection	75
5.2	Nonlinear isotropic diffusion filtering	82
6	Conclusions	87
A	Definitions and fundamental theorems	89
	List of notations	93
	Bibliography	102

Chapter 1

Introduction

When one looks at a digital image, it is impossible to avoid seeing in them structures, which in a fraction of a second, can be identified with real objects. To teach a computer to carry out this task, which is easy for a human, is one of the most difficult problems of image processing and analysis. Foremost, the only information which computer gets is the intensity values at pixels, and the whole processing of an image is usually divided into several stages: filtering, segmentation, classification and finally recognition.

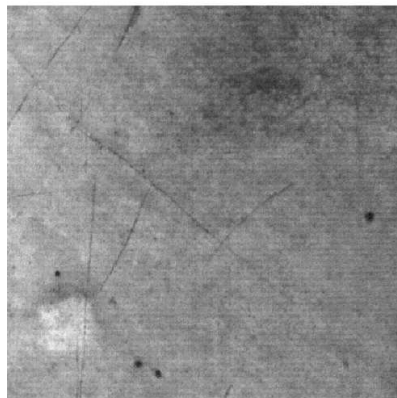


Fig. 1.1: The image of a leather with defects

In order to explain each of these individual stages, as an example, let us consider the real problem¹ concerning detection and recognition of defects on a leather. In Fig. 1.1, one can distinguish various imperfections such as

¹The problem comes from the Image Processing Department of the Institute for Industrial Mathematics in Kaiserslautern

scratches (straight lines), bites (black dots) and shingle (light stain). The first step in the processing of such image is filtering. The aim of this technique is not only to remove a noise and smoothening parts of leather, where it is inhomogeneous because of its complex structure, but also to enhance details of interest that are obscured. Shortly speaking, filtering should be performed in such a way to simplify the problem of segmentation. Segmenting an image consists of finding its meaningful regions or its edges. With reference to the mentioned example, the meaningful regions should correspond to defects and edges to their apparent contours. Once all defects are detected, we have to classify them. It means, we have to divide the set of these defects based on its attributes (straight, black, light, round, etc.) into three classes. Finally, the problem of the recognition is reduced to assigning the respective name to each of the class. Therefore, at the final result, the important influence has all previous stages of processing and the approach chosen to solve the problem on each of them, should simplify the problem on the next one.

In this thesis we consider theoretically and numerically the first problem of the above described example. This refers to the problem of filtering with preserving and enhancing edges. Nowadays, there are a large number of methods originated in many mathematical theories, which have been proposed for this problem. From stochastic and statistical modeling [14, 26, 34, 36] through signal processing techniques, including wavelets [28, 27] and other transform theories to approaches based on partial differential equations [62, 48, 10]. We refer to the book of Chan and Shen [22] for complete and informative review of all mentioned approaches as well as to the paper of Mrázek et al. [50], where authors deal with establishing relations between number of widely used nonlinear filters in image processing.

In the further part of thesis, we focus only on techniques based on partial differential equations (PDEs). The starting point to first mathematical justifications of this approaches, was the simple observation of Koenderink [41] that commonly used as a filter, convolution of an image with the Gaussian kernel at each scale is equivalent to the solution of linear diffusion problem with this image as an initial condition. The next, an important step in development of this theory was introduction by Perona and Malik in [52] nonlinear diffusion model with a more accurate edge detection. In this publication authors, as a first, state explicitly a maximum principle as a basic requirement in image processing, cf. [2]. The Perona and Malik model was later axiomatized [2], regularized [18] and modified [3] by Lions and Morel et al. The important contribution to diffusion filter theory is also the work of Weickert [63], where author proposed an extension of the model presented in [18] to the anisotropic case.

Let us now explain the main idea of the nonlinear diffusion filtering. For that, first we have to define what do we mean by a digital image. From the mathematical point of view, digital image is a bounded function f defined on domain Ω in \mathbb{R}^2 , assigning a value $f(x)$ to each point x in this domain. In the case of gray images this value is a real number and in the case of colour images, it is a real valued vector with three components. The main idea of the nonlinear diffusion filtering is based upon adaptation of the diffusion process to the image structure. Such an approach allows us to preserve or even enhance edges and at the same time to smooth regions of an image, which do not contain any important information that could be used in the next stages of processing. The main problem, when we want to define a nonlinear filter is to know a position of objects boundary. Of course, this information is not known in advance and the best what we can do is to estimate the boundary location by some edge detector. This means, we have to do preliminary step toward segmentation. In the introduction of the book [49], Morel and Solimini write

(...) most segmentation algorithms try to minimize, by several very different procedures, one and the same segmentation energy. This energy measures how smooth the regions are, how faithful the analyzed image to the original image and the obtained edges to the image discontinuities are.

In the continuous setting this energy has been defined by Mumford and Shah in [51], as follows

$$J(u, E) := \frac{1}{2} \int_{\Omega} (u - f)^2 dx + \frac{1}{2} \int_{\Omega \setminus E} |\nabla u|^2 dx + \mathcal{H}^1(E) \quad (1.1)$$

where E denotes the set of edges and $\mathcal{H}^1(E)$ its one-dimensional Hausdorff measure. Let us now assume that the set E_1 better approximates the set of edges in image f than E_2 , and that the function u minimizes functional J . Then, obviously we have

$$J(u, E_1) \leq J(u, E_2)$$

Based on this idea, our approach proposes to consider the following functional

$$J_{\varepsilon}(u) := \frac{1}{2} \int_{\Omega} (u - f)^2 dx + \frac{1}{2} \int_{\Omega} \alpha_{\varepsilon} |\nabla u|^2 dx \quad (1.2)$$

with the piecewise constant function α_ε , defined by

$$\alpha_\varepsilon(x) := \begin{cases} \alpha & \text{if } x \in B_\varepsilon^j, \ j = 1, \dots, m \\ 1 & \text{if } x \in \Omega \setminus \bigcup_{j=1}^m \bar{B}_\varepsilon^j \end{cases}$$

In the above definition $\alpha \ll 1$ is a positive constant and B_ε^j , for $j = 1, \dots, m$, are small inhomogeneities, centred at points belonging to the set of edges E . Following the above reasoning, we can write

$$J_\varepsilon(u) \leq J_0(u) \tag{1.3}$$

where functional J_0 is defined by

$$J_0(u) := \frac{1}{2} \int_{\Omega} (u - f)^2 dx + \frac{1}{2} \int_{\Omega} |\nabla u|^2 dx$$

Let us assume that function u_ε is a global minimizer of functional J_ε . Then, we have $J_\varepsilon(u_\varepsilon) \leq J_\varepsilon(v)$ for any function v . Using this fact and inequality (1.3), we obtain

$$J_\varepsilon(u_\varepsilon) \leq J_0(u_0)$$

where function u_0 , taken as v , is a global minimizer of functional J_0 . This observation gives us motivation to investigate an asymptotic behavior of the following difference

$$J_\varepsilon(u_\varepsilon) - J_0(u_0) \tag{1.4}$$

as ε tends to 0. The main goal of this thesis is to find an explicit formula for a dominant term in an asymptotic expansion of expression (1.4). The theoretical background of our approach is based on the idea of the topological sensitivity analysis. In the classical formulation, this theory, provides a variation of a given functional J with respect to subtraction from the domain Ω a small ball B_ε . This variation is a scalar function, independent on ε , called the topological derivative or the topological gradient and is used as a descent direction to solve various problems, among other things topological shape optimization [59, 1, 20], inverse problems [32], and recently also image segmentation and restoration [45, 11]. For the first time, the definition of topological gradient has been introduced in the context of compliance optimization for linear elasticity problems by Schumacher in [57]. The first

mathematical justification of this method has been done by Sokołowski et al. in [58, 59]. In [35], using an adaptation of the adjoint method [19] and a domain truncation technique, Garreau et al. presented a method to obtain the topological asymptotic expansion. This approach was later investigated also by Feijóo et al. in [33]. In [9, 7] Amstutz et al. modified the definition of the topological gradient and have proposed to provide a variation of a functional with respect to a change of some material properties. Recent results concerning an incorporation of the topological derivative into level set method are presented in [8, 16]. In this thesis, we based on ideas presented by Garreau et al. in [35] and by Amstutz et al. in [6, 9].

Since PDEs are written in the continuous settings, once the filter model is defined, we have to discretize it in order to find a numerical solution. For this purpose, several kinds of approaches can be considered. In the image processing community the finite difference method is the most popular one. There are many publications where the continuous model of nonlinear diffusion filters [18, 63, 64], are discretized using this technique. However, as the structure of digital image is a set of uniformly distributed pixels, the approximation on cell-centred grid using the finite volume method, seems to be a natural choice for image processing applications, mainly for the sake of a more clear description of an image boundary and thereby treatment of the boundary condition. Another advantage of this method is its easy implementation along with the possibility of discretizing the problem on a nonuniform grid, adapted to the local structure of an image. Such an approach has practical importance in the case of solving real problems like the one described in the beginning of this introduction. Usually, in solving such problems we have to deal with images of large sizes, while regions containing important information often take only a few percent of the whole image. In order to substantially save computer storage during processing of such images one can apply the grid coarsening technique which allows to reduce the size of data rapidly. Among the many publications that consider the finite volume discretization of the PDEs on uniform and nonuniform [30, 31] grids, there are only a few which are suggested for image processing applications. The numerical solution of the Catté et al. model [18] using the semi-implicit finite volume approximation scheme and the proof of its convergence was proposed by Mikula and Ramarosy in [47]. In papers [44] and [43] the coarsening strategy for the computational method has been presented. The adaptivity for the finite element method in image processing applications has been suggested in [12] and generalized to the three dimensional case in [13]. The approach given in [12] has been modified by Preusser and Rumpf in [54]. In this paper, authors improve and discuss storage requirements for the method.

The thesis is organized as follows. In Chapter 2, we give a brief overview of the well known PDE-based methods to image filtering and explain some its theoretical and practical difficulties. The main result of this thesis is contained in Chapter 3. In the first section of this chapter, we formulate mathematical problem and prove its well-posedness. Next part is devoted to asymptotic analysis of expression (1.4). In Section 3.2 of Chapter 3, based on results of Vogelius et al. [61], we derive an asymptotic formula for $u_\varepsilon - u_0$ on the boundary of inhomogeneity, which we apply in the proof of Proposition 3.1. In Section 3.3 of Chapter 3, explicit formulas of the topological gradient for two different shapes of inhomogeneities are presented. At the end of Chapter 3, we give some remarks and discuss problems which appear, when we try to apply some of existing method of the topological sensitivity analysis to our particular problem. In Chapter 4, we propose discretization of the Catté [18] and the Weickert [63] model based on finite volume technique using the integro-interpolation method introduced by Samarskii in [56]. Proposed discretization is derived for the case of the uniform as well as for the case of nonuniform cell-centered grid obtained by application of the adaptive coarsening strategy. Numerical experiments are presented in Chapter 5. At the end some conclusions are given. In Appendix A, we list the fundamental theorems and definitions used in the thesis.

Chapter 2

An overview of diffusion filters

The aim of this chapter is to present some classical PDE-based method for image filtering and discuss its practical and theoretical difficulties. In this thesis, for a simplicity of the presentation, we will consider only gray images. This means, that we will assume that the initial image f is a real function in class $L^\infty(\Omega)$, defined on an open and bounded domain $\Omega \subset \mathbb{R}^2$. Values $f(x)$ represent brightness or gray level of the image at each point $x \in \Omega$.

2.1 Linear diffusion filter

As we have already mentioned, the oldest and most investigated PDE in image processing is the parabolic linear diffusion equation of the form

$$\frac{\partial u}{\partial t}(x, t) = \Delta u(x, t) \quad (x, t) \in \mathbb{R}^2 \times (0, \infty) \quad (2.1)$$

with the initial condition $u(x, 0) = f(x)$ for any $x \in \mathbb{R}^2$. Note that we have here $x \in \mathbb{R}^2$. In fact, we consider that f is primarily defined only on the domain Ω . Nevertheless, by symmetry and then periodicity we can extend it to \mathbb{R}^2 . This method of extension is typical in image processing.

The underlying idea to apply equation (2.1) in image processing comes from the early work of Koenderink [41], who noticed that widely used in denoising, convolution of the image f with the Gaussian kernel, defined by

$$G_\sigma(x, y) := \frac{1}{2\pi\sigma^2} \exp\left(-\frac{|x - y|^2}{2\sigma^2}\right) \quad (2.2)$$

is equivalent to the solution u of problem (2.1) for $t = \frac{1}{2}\sigma^2$, that is

$$u(x, t) = (G_{\sqrt{2t}} * f)(x) = \int_{\mathbb{R}^2} G_{\sqrt{2t}}(x, y) f(y) dy$$

The above formula gives the correspondence between the time t and the scale parameter σ of the Gaussian kernel G_σ .

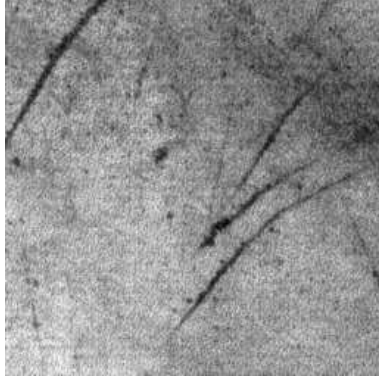


Fig. 2.1: Input image $f \in [0, 255]$, $\Omega = [0, 256]^2$

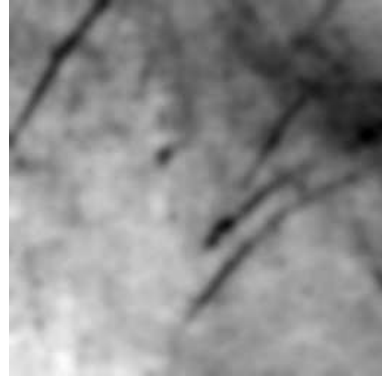


Fig. 2.2: Application of the linear diffusion filter with parameters $\tau = 0.5$, $iter = 10$

The linear diffusion filter has one serious disadvantage. As a matter of fact, it smoothens an image but at the same time blurs important features such as edges (see Fig. 2.1 and Fig. 2.2) making it difficult to identify on the next stage of image analysis, namely segmentation. To overcome this problem one should consider a nonlinear filter, adapted to the local image structure.

2.2 Nonlinear isotropic diffusion filter

2.2.1 The Perona and Malik model

For the first time a nonlinear diffusion filter was introduced by Perona and Malik in [52]. They proposed to replace (2.1) by a nonlinear diffusion equation with homogeneous Neumann conditions on the boundary $\partial\Omega$ and to solve the following problem in order to obtain the smoothed version of the initial image f

$$\begin{cases} \frac{\partial u}{\partial t} = \nabla \cdot (d(|\nabla u|^2) \nabla u) & (x, t) \in \Omega \times (0, T] \\ u(x, 0) = f & x \in \Omega \\ \frac{\partial u}{\partial n} = 0 & (x, t) \in \Omega \times (0, T] \end{cases} \quad (2.3)$$

In the first equation, the diffusivity d is a positive, monotonically decreasing function, defined in a way, such that the smoothing of image is conditional and depends on its structure. If $|\nabla u|^2$ is large, then the diffusion is low and therefore the exact location of the edges is kept. If $|\nabla u|^2$ is small, then the diffusion tends to smooth more around x . Of course, there exist several possible choices for the diffusivity d . As an example, authors proposed to consider the following definition

$$d(s) := \frac{1}{1 + s/\mu} \quad (2.4)$$

where the parameter $\mu > 0$ plays the role of a threshold. The plot of this function is presented in Fig. 2.3.

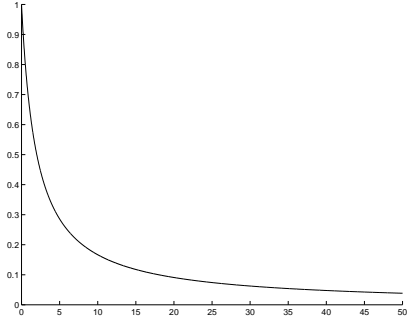


Fig. 2.3: Plot of the diffusivity $d(s) = \frac{1}{1+s/\mu}$ with a threshold $\mu = 2$

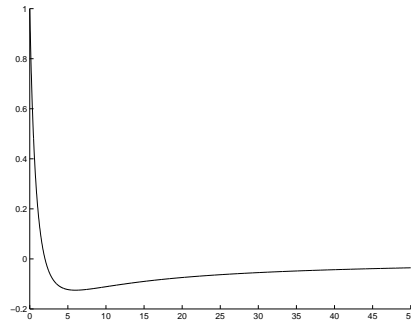


Fig. 2.4: Plot of the function $b(s) = d(s) + 2s d'(s)$

However, the Perona and Malik model has several practical and theoretical difficulties. If the image is noisy, then the noise would introduce very large oscillations of the gradient ∇u . Therefore, the adaptive smoothing introduced by the model (2.3) would not give good results, since all these noise edges will be kept. The second difficulty arises from the fact that we obtain a backward diffusion equation for $|\nabla u|^2 > \mu$, which is a classical example of the ill-posed problem. In practice, it implies that very similar images can give divergent solutions and therefore different edges.

Let us explain the second problem in detail. For that, let $\eta = \nabla u / |\nabla u|$ and $\nu = \nabla u^\perp / |\nabla u|$ be a vector parallel and perpendicular, respectively, to the gradient ∇u and let us decompose the divergence operator in (2.3) using directions η and ν . We have

$$\nabla \cdot (d(|\nabla u|^2) \nabla u) = d(|\nabla u|^2) \Delta u + 2 d'(|\nabla u|^2) \langle \nabla u, \nabla^2 u \nabla u \rangle \quad (2.5)$$

where the expression $\langle \nabla u, \nabla^2 u \nabla u \rangle$ is nothing but the second order derivative of the function u in the gradient ∇u direction.

From the other side, we have

$$\frac{\partial^2 u}{\partial \eta^2} = \langle \eta, \nabla^2 u \eta \rangle = \frac{1}{|\nabla u|^2} \left[\left(\frac{\partial u}{\partial x_1} \right)^2 \frac{\partial^2 u}{\partial x_1^2} + 2 \frac{\partial u}{\partial x_1} \frac{\partial u}{\partial x_2} \frac{\partial^2 u}{\partial x_1 \partial x_2} + \left(\frac{\partial u}{\partial x_2} \right)^2 \frac{\partial^2 u}{\partial x_2^2} \right]$$

and

$$\frac{\partial^2 u}{\partial \nu^2} = \langle \nu, \nabla^2 u \nu \rangle = \frac{1}{|\nabla u|^2} \left[\left(\frac{\partial u}{\partial x_2} \right)^2 \frac{\partial^2 u}{\partial x_1^2} - 2 \frac{\partial u}{\partial x_1} \frac{\partial u}{\partial x_2} \frac{\partial^2 u}{\partial x_1 \partial x_2} + \left(\frac{\partial u}{\partial x_1} \right)^2 \frac{\partial^2 u}{\partial x_2^2} \right]$$

Thus, the first equation in (2.3) may be written as follows

$$\frac{\partial u}{\partial t} = d(|\nabla u|^2) \frac{\partial^2 u}{\partial \nu^2} + b(|\nabla u|^2) \frac{\partial^2 u}{\partial \eta^2} \quad (2.6)$$

where $b(s) = d(s) + 2sd'(s)$. Thus, we can interpret equation (2.6) as the sum of a diffusion in the η and ν directions, with functions d and b acting as weighting coefficients.

Let us first check the parabolicity of equation (2.6) for an arbitrary function d . From Evans [29, page 350], we have the following definition

Definition 2.1 *The partial differential operator $\frac{\partial}{\partial t} + L$, where*

$$Lu = - \sum_{i,j=1}^N a_{ij}(x,t) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^N b_i(x,t) \frac{\partial u}{\partial x_i} + c(x,y)u$$

is parabolic for given coefficients a_{ij} , b_i , c ($i, j = 1, \dots, N$) if and only if there exist constant $C > 0$, such that

$$\sum_{i,j=1}^N a_{ij}(x, t) \xi_i \xi_j \geq C |\xi|^2$$

for all $(x, t) \in \Omega \times (0, T]$ and $\xi \in \mathbb{R}^N$.

We observe that equation (2.6) may be written as follows

$$\frac{\partial u}{\partial t} = a_{11}(|\nabla u|^2) \frac{\partial^2 u}{\partial x_1^2} + 2 a_{12}(|\nabla u|^2) \frac{\partial^2 u}{\partial x_1 \partial x_2} + a_{22}(|\nabla u|^2) \frac{\partial^2 u}{\partial x_2^2} \quad (2.7)$$

with

$$\begin{aligned} a_{11}(|\nabla u|^2) &= 2 \left(\frac{\partial u}{\partial x_1} \right)^2 d'(|\nabla u|^2) + d(|\nabla u|^2) \\ a_{12}(|\nabla u|^2) &= 2 \frac{\partial u}{\partial x_1} \frac{\partial u}{\partial x_2} d'(|\nabla u|^2) \\ a_{22}(|\nabla u|^2) &= 2 \left(\frac{\partial u}{\partial x_2} \right)^2 d'(|\nabla u|^2) + d(|\nabla u|^2) \end{aligned}$$

According to Definition 2.1, equation (2.7) is parabolic if and only if

$$\sum_{i=1,2} a_{ij}(|\nabla u(x, t)|^2) \xi_i \xi_j > 0$$

for all $(x, t) \in \Omega \times (0, T]$ and all $\xi \in \mathbb{R}^2$.

An easy calculation shows that this condition reduces to the single inequality

$$b(s) > 0$$

Let us now examine problem (2.3) with the diffusivity d defined as in (2.4). We have

$$d'(s) = \frac{-1}{\mu(1 + s/\mu)^2}$$

and

$$b(s) = d(s) + 2sd'(s) = \frac{\mu(\mu - s)}{(\mu + s)^2}$$

Therefore, we get $b(|\nabla u|^2) \leq 0$ for $|\nabla u|^2 \geq \mu$. This implies that the model (2.3) with diffusivity d defined in (2.4) fulfill our expectations: it is a backward in the direction perpendicular to ∇u allowing us to sharpen edges. However a unique solution to problem (2.3) does not exist.

2.2.2 Regularization of the Perona and Malik model

One way to deal with an ill-posed problem like (2.3) is to introduce regularization, which would make the problem well-posed. Then, by reducing the amount of regularization and observing the behavior of the solution of the regularized problem, we can obtain information about the initial one. In the first time, the method to regularize the Perona and Malik problem was proposed by Catté et al. in [18]. To avoid both of mentioned in the previous subsection problems, they suggested to replace the gradient ∇u in the diffusivity $d(|\nabla u|^2)$ by its estimate $\nabla u_\sigma := \nabla G_\sigma * u$, where G_σ is the Gaussian kernel as defined in (2.2). They have also proven that this slight change is sufficient to ensure existence and uniqueness of the solution to the problem (2.3). This result gives the following theorem

Theorem 2.1 *Let $d : \mathbb{R}_+ \cup \{0\} \rightarrow \mathbb{R}_+$ be smooth, decreasing with $d(0) = 1$, $\lim_{s \rightarrow \infty} d(s) = 0$. If $f \in L^2(\Omega)$, then there exists a unique function $u(x, t) \in C([0, T]; L^2(\Omega)) \cap L^2((0, T); H^1(\Omega))$ satisfying in the distributional sense*

$$\begin{cases} \frac{\partial u}{\partial t} = \nabla \cdot (d(|\nabla u_\sigma|^2) \nabla u) & (x, t) \in \Omega \times (0, T] \\ u(x, 0) = f & x \in \Omega \\ \frac{\partial u}{\partial n} = 0 & (x, t) \in \partial\Omega \times (0, T] \end{cases} \quad (2.8)$$

Moreover, $|u|_{L^\infty((0, T); L^2(\Omega))} \leq |f|_{L^2(\Omega)}$ and $u \in C^\infty((0, T) \times \bar{\Omega})$

Proof.

Here we give only the main idea of the proof. For the complete proof we refer to [18].

To prove uniqueness, we use energy estimates for the difference of two solutions to problem (2.8), so that the Gronwall inequality can be applied.

Then, uniqueness follows from the fact that both solutions start with the same initial values.

The proof of existence is based on the Schauder fixed point theorem. We consider a variational version of the problem (2.8) with diffusivity d depending on some function w instead of u . This problem is now linear in u and has a unique solution u_w in space $W(0, T)$, defined by

$$W(0, T) := \left\{ w \in L^2((0, T); H^1(\Omega)), \frac{dw}{dt} \in L^2((0, T); H^1(\Omega)') \right\}$$

Next, we can prove that $w \rightarrow S(w) \equiv u_w$ is a weakly continuous mapping from a non empty, convex and weakly compact subset W_0 of $W(0, T)$ into itself. Owing to the Schauder fixed point theorem, there exists $w \in W_0$ such that $w = S(w) = u_w$. Thus, the function u_w solves the problem (2.8).

The regularity follows from the general theory of parabolic equations

□

Another recent approach to regularize the Perona and Malik equation has been introduced by Amann in [4]. Instead of the above mentioned space regularization, he proposed to consider regularization in time. Analytical results presented in his paper based on a new sharp existence and uniqueness theorem for quasilinear parabolic evolution equations.

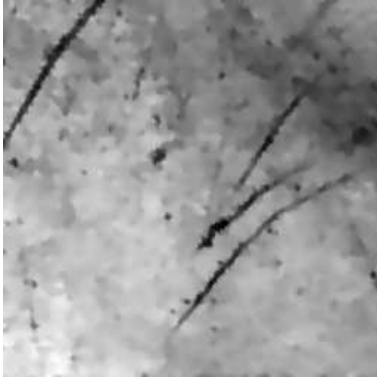


Fig. 2.5: Application of the nonlinear isotropic diffusion filter with parameters $\mu = 20$, $\sigma = 1$, $\tau = 0.5$, $iter = 10$

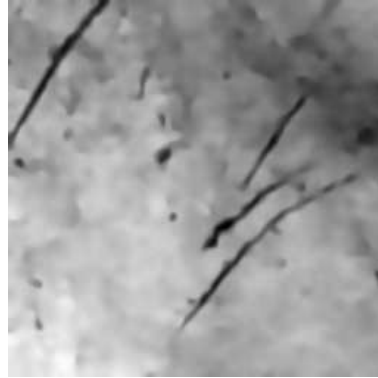


Fig. 2.6: Application of the nonlinear anisotropic diffusion filter with parameters $\mu = 20$, $\sigma = 1$, $\rho = 1$, $\tau = 0.5$, $iter = 10$

2.3 Nonlinear anisotropic diffusion filter

2.3.1 The Weickert model

Despite of the advantages of the isotropic diffusion filter, there is still one imperfection: when the diffusion process is stopped near the boundary of an object, it preserves the edges but also leaves a noise at these positions (see Fig. 2.5). To avoid this effect, Weickert in [63] suggested to modify the diffusion operator so that it diffuses more in direction parallel to edges and less in the perpendicular one. In order to filter an image, he proposed to consider the following problem

$$\begin{cases} \frac{\partial u}{\partial t} = \nabla \cdot (D(S_\rho(\nabla u_\sigma)) \nabla u) & (x, t) \in \Omega \times (0, T] \\ u(x, 0) = f & x \in \Omega \\ \langle D(S_\rho(\nabla u_\sigma)) \nabla u, n \rangle = 0 & (x, t) \in \partial\Omega \times (0, T] \end{cases} \quad (2.9)$$

where $D(S_\rho(\nabla u_\sigma))$ is symmetric, positive semidefinite matrix, called the diffusion tensor and it is constructed in the way described below.

To avoid false detections of edges due to the presence of noise, we first convolve u with the Gaussian kernel G_σ and calculate the matrix

$$S_0(\nabla u_\sigma) := \nabla u_\sigma^T \nabla u_\sigma \quad (2.10)$$

This matrix possesses an orthogonal basis composed of eigenvectors v_1, v_2 with $v_1 \parallel \nabla u_\sigma$ and $v_2 \perp \nabla u_\sigma$. The corresponding eigenvalues are equal to $|\nabla u_\sigma|^2$ and 0, respectively, and give contrast in direction of eigenvectors.

In the next step, the local information is averaged by convolving S_0 componentwise with the Gaussian kernel G_ρ . As a result we obtain symmetric, positive semidefinite matrix

$$S_\rho(\nabla u_\sigma) := G_\rho * S_0(\nabla u_\sigma) := \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} \quad (2.11)$$

The matrix S_ρ is called the structure tensor and possesses orthonormal eigenvectors v_1, v_2 with v_1 parallel to

$$\begin{bmatrix} 2s_{12} \\ s_{22} - s_{11} + \sqrt{(s_{11} - s_{22})^2 + 4s_{12}^2} \end{bmatrix}$$

The corresponding eigenvalues are given by

$$\mu_1 = \frac{1}{2} \left[s_{11} + s_{22} + \sqrt{(s_{11} - s_{22})^2 + 4s_{12}^2} \right]$$

and

$$\mu_2 = \frac{1}{2} \left[s_{11} + s_{22} - \sqrt{(s_{11} - s_{22})^2 + 4s_{12}^2} \right]$$

The diffusion tensor $D(S_\rho(\nabla u_\sigma))$ should have the same set of eigenvectors as the structure tensor S_ρ , and the choice of corresponding eigenvalues should depend on the desired goal of the filter. Further, we will introduce two possible choices of eigenvalues λ_1 and λ_2 of $D(S_\rho(\nabla u_\sigma))$

In [63], Weickert has presented the following result

Theorem 2.2 *Let us assume that:*

- (i) *The diffusion tensor $D(S_\rho(\nabla u_\sigma))$ belongs to $C^\infty(\mathbb{S}^{2 \times 2})$*
- (ii) *For all $w \in L^2(\Omega, \mathbb{R}^2)$ with $|w(x)| \leq k$ on $\bar{\Omega}$, there exists a positive lower bound $\nu(k)$ for the eigenvalues of $D(S_\rho(w))$*

Then for all $f \in L^\infty(\Omega)$ problem (2.9) has unique solution $u(x, t)$ satisfying

$$u \in C([0, T]; L^2(\Omega)) \cap L^2([0, T]; H^1(\Omega))$$

$$\frac{\partial u}{\partial t} \in L^2([0, T]; H^1(\Omega))$$

Moreover, $u \in C^\infty((0, T) \times \bar{\Omega})$. This solution depends continuously on f with respect to the L^2 -norm, and it satisfies the extremum principle

$$\inf_{\Omega} f(x) \leq u(x, t) \leq \sup_{\Omega} f(x)$$

Proof.

Existence, uniqueness and regularity are straightforward anisotropic extensions of the proof for the isotropic case (see proof of Theorem 2.1). The proof of the maximum and minimum principle is based on the Stampacchia truncation method. For the complete proof we refer to [63]

□

Related results have been proved for semi-discrete and fully discrete version of the model. We refer to Weickert [63] for further proofs of properties (invariances, image simplification properties, behavior as t tends to infinity).

2.3.1.1 Definition of the diffusion tensor

Let us now describe how to define the diffusion tensor $D(S_\rho(\nabla u_\sigma))$ such that the anisotropic filter (2.9) could be applied to particular problems.

Since the eigenvectors of the diffusion tensor should reflect the local image structure, one should choose the same orthonormal basis of eigenvectors that one gets from the structure tensor S_ρ . Following Weickert [63], we introduce here two possible choices of eigenvalues λ_1 and λ_2 of $D(S_\rho(\nabla u_\sigma))$

Edge-enhancing anisotropic diffusion. If one wants to smooth an image within some region and wants to preserve edges, then one should reduce the diffusivity λ_1 perpendicular to edges more if the contrast given by the greatest eigenvalue μ_1 of the structure tensor S_ρ is large. This behavior may be accomplished by the following choice of eigenvalues

$$\lambda_1 := \begin{cases} 1 & \text{if } \mu_1 \leq 0 \\ 1 - \exp\left(\frac{-3.315}{\mu_1^4}\right) & \text{if } \mu_1 > 0 \end{cases}$$

$$\lambda_2 := 1$$

Coherence-enhancing anisotropic diffusion. If one wants to enhance coherent structures, then one should perform smoothening, preferably along the coherence direction v_2 with diffusivity λ_2 , which is increasing with respect to the coherence $(\mu_1 - \mu_2)^2$. This may be achieved by the following choice for the eigenvalues of the diffusion tensor

$$\lambda_1 := \alpha$$

$$\lambda_2 := \begin{cases} \alpha & \text{if } \mu_1 = \mu_2 \\ \alpha + (1 - \alpha) \exp\left(\frac{-1}{(\mu_1 - \mu_2)^2}\right) & \text{otherwise} \end{cases}$$

where $\alpha \in (0, 1)$ is a small positive parameter which keeps the diffusion tensor $D(S_\rho(\nabla u_\sigma))$ uniformly positive definite.

Chapter 3

Topological sensitivity analysis

In this chapter, based on the idea presented by Amstutz et al. in [7, 9], we investigate an asymptotic behavior of the cost function J_ε , defined as in (1.2), with respect to the change of diffusivity¹. Unlike Amstutz et al. we provide an asymptotic expansion of the form

$$J_\varepsilon(u_\varepsilon) - J_0(u_0) = \varepsilon^2 g(x) + o(\varepsilon^2) \quad (3.1)$$

where functions u_ε and u_0 are global minimizers of the functionals J_ε and J_0 , respectively. This means

$$J_\varepsilon(u_\varepsilon) = \min_{u \in V} J_\varepsilon(u) \quad \text{and} \quad J_0(u_0) = \min_{u \in V} J_0(u)$$

The essential contribution for this chapter is Theorem 3.4 proposed and proved by Vogelius and Volkov in [61]. In this publication, authors provide asymptotic formulas for perturbations in the electromagnetic fields due to the presence of small inhomogeneities. The function u_ε , being the asymptotic behavior they consider is a solution of the problem

$$\begin{cases} \nabla \cdot \left(\frac{1}{\mu_\varepsilon} \nabla u_\varepsilon \right) + \omega^2 \left(\epsilon_\varepsilon + i \frac{\sigma_\varepsilon}{\omega} \right) u_\varepsilon = 0 & x \in \Omega \\ u_\varepsilon = f & x \in \partial\Omega \end{cases} \quad (3.2)$$

where $\omega > 0$ is a given frequency. The electric permability is defined as follows

¹In the sequel, we will use the term ‘topological gradient’ or ‘topological derivative’ despite we do not change the topology of the domain Ω .

$$\mu_\varepsilon(x) := \begin{cases} \mu_j & \text{if } x \in B_\varepsilon^j, \ j = 1, \dots, m \\ \mu_0 & \text{if } x \in \Omega \setminus \bigcup_{j=1}^m \bar{B}_\varepsilon^j \end{cases} \quad (3.3)$$

In the above definition, the numbers μ_0 and μ_j , for $j = 1, \dots, m$, are constant. The electric permittivity ϵ_ε and electric conductivity σ_ε in the first equation of problem (3.2) are defined analogously.

3.1 Formulation of the problem

Let Ω be a bounded open domain in \mathbb{R}^2 with C^1 boundary $\partial\Omega$, in particular, Ω can be consider as a rectangle. Our goal is to prove that for sufficiently small $\varepsilon \geq 0$, there exists a function g , independent on ε , such that the following topological asymptotic expansion holds

$$J_\varepsilon(u_\varepsilon) - J_0(u_0) = \varepsilon^2 g(x) + o(\varepsilon^2)$$

Let us recall that the functional $J_\varepsilon : H^1(\Omega) \rightarrow \mathbb{R}$ in the above expression, is defined by

$$J_\varepsilon(u) := \frac{1}{2} \int_{\Omega} (u - f)^2 dx + \frac{1}{2} \int_{\Omega} \alpha_\varepsilon |\nabla u|^2 dx \quad (3.4)$$

where the image f is a given function in $L^\infty(\Omega)$ and piecewise constant diffusivity α_ε is defined as follows

$$\alpha_\varepsilon(x) := \begin{cases} \alpha & \text{if } x \in B_\varepsilon^j, \ j = 1, \dots, m \\ 1 & \text{if } x \in \Omega \setminus \bigcup_{j=1}^m \bar{B}_\varepsilon^j \end{cases} \quad (3.5)$$

The positive constant number $\alpha \ll 1$ is a regularization parameter and $B_\varepsilon^j = x_j + \varepsilon B$, for $j = 1, \dots, m$, are small inhomogeneities centred at the point $x_j \in \Omega$. The fixed, bounded open domain $B \subset \Omega$ describe their relative shape.

The functions u_ε and u_0 are minimisers of functionals J_ε and J_0 , respectively, where J_0 is defined by

$$J_0(u) := \frac{1}{2} \int_{\Omega} (u - f)^2 dx + \frac{1}{2} \int_{\Omega} |\nabla u|^2 dx \quad (3.6)$$

Let us remark that the cost function J_ε is defined similarly to the Mumford-Shah functional (1.1), introduced in the context of variational approach to image segmentation in [51] and investigated in the book of Morel and Solimini [49]. The first term in (3.91) ensures that the minimizer of the functional J_ε is close to the input image f and the second term imposes to smooth u only within homogeneous regions.

Let us consider the minimization problem

$$\text{Find } u \in H^1(\Omega) \text{ such that } J_\varepsilon(u) = \min_{v \in H^1(\Omega)} J_\varepsilon(v) \quad (3.7)$$

We will show that for fixed α_ε , the minimum of the functional J_ε is achieved at $u_\varepsilon \in H^1(\Omega)$, solution of the problem

$$\begin{cases} u_\varepsilon - \nabla \cdot (\alpha_\varepsilon \nabla u_\varepsilon) = f & x \in \Omega \\ \frac{\partial u_\varepsilon}{\partial n} = 0 & x \in \partial\Omega \end{cases} \quad (3.8)$$

The function $u_\varepsilon \in H^1(\Omega)$ on the boundary ∂B_ε^j satisfies

$$\begin{cases} u_\varepsilon^+ - u_\varepsilon^- = 0 & x \in \partial B_\varepsilon^j \\ \frac{\partial u_\varepsilon^+}{\partial n} - \alpha \frac{\partial u_\varepsilon^-}{\partial n} = 0 & x \in \partial B_\varepsilon^j \end{cases} \quad (3.9)$$

for $j = 1, \dots, m$.

The weak formulation associated with problem (3.8) is as follows

$$\begin{cases} \text{Find } u \in H^1(\Omega) \\ a_\varepsilon(u, v) = l(v) \quad \forall v \in H^1(\Omega) \end{cases} \quad (3.10)$$

where the bilinear form $a_\varepsilon : H^1(\Omega) \times H^1(\Omega) \rightarrow \mathbb{R}$ is defined by

$$a_\varepsilon(u, v) := \int_{\Omega} uv \, dx + \int_{\Omega} \alpha_\varepsilon \nabla u \cdot \nabla v \, dx \quad (3.11)$$

and the linear form $l : H^1(\Omega) \rightarrow \mathbb{R}$ by

$$l(v) := \int_{\Omega} f v \, dx \quad (3.12)$$

To prove an equivalence of the three above mentioned problems we will make use of following lemmas

Lemma 3.1 *Let $v \in L^1_{loc}(\Omega)$, where Ω is a domain in \mathbb{R}^N and let*

$$\int_{\Omega} v \varphi \, dx = 0$$

holds for any function $\varphi \in C_0^\infty(\Omega)$. Then $v \equiv 0$ almost everywhere.

Lemma 3.2 *Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with C^1 boundary $\partial\Omega$. Then $C_0^\infty(\Omega)$ is a dense subset of $H^1(\Omega)$.*

Lemma 3.1 is called the fundamental lemma of calculus of variations and the proof can be found in [39, page 6]. The proof of Lemma 3.2 can be found in [38, page 117].

Theorem 3.1 *The three problems (3.7), (3.8) and (3.10) are equivalent.*

Proof.

(3.8) \Rightarrow (3.10)

Suppose $u \in H^1(\Omega)$ is a solution of problem (3.8). Let us multiply the first equation in (3.8) by an arbitrary test function $\varphi \in C_0^\infty(\Omega)$ and then apply the Gauss theorem. We have

$$\int_{\Omega} u \varphi \, dx + \int_{\Omega} \alpha_\varepsilon \nabla u \cdot \nabla \varphi \, dx = \int_{\Omega} f \varphi \, dx \quad \forall \varphi \in C_0^\infty(\Omega) \quad (3.13)$$

From Lemma 3.2 we know that $C_0^\infty(\Omega)$ is dense in $H^1(\Omega)$, which means that for any $v \in H^1(\Omega)$ there exists a sequence $\varphi_n \in C_0^\infty(\Omega)$ such that $\lim_{n \rightarrow \infty} \varphi_n = v$ in $H^1(\Omega)$. This implies

$$\int_{\Omega} uv \, dx + \int_{\Omega} \alpha_\varepsilon \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx \quad \forall v \in H^1(\Omega) \quad (3.14)$$

(3.10) \Rightarrow (3.8)

Suppose now $u \in H^1(\Omega)$ is a solution of problem (3.10). By taking $v \in C_0^\infty(\Omega) \subset H^1(\Omega)$ and applying integration by parts formula we obtain

$$\int_{\Omega} [u - f - \nabla \cdot (\alpha_\varepsilon \nabla u)] v \, dx + \int_{\partial\Omega} \alpha_\varepsilon \frac{\partial u}{\partial n} v \, ds = 0 \quad (3.15)$$

Using Lemma 3.1 and assuming $v \in C_0^\infty(\Omega)$ is an arbitrary function, we obtain the desired result.

(3.10) \Rightarrow (3.7)

Let us write the functional (3.91) with the help of bilinear (3.11) and linear (3.12) form as follow

$$J_\varepsilon(u) = \frac{1}{2}a_\varepsilon(u, u) - l(u) + \frac{1}{2} \int_\Omega f^2 dx$$

and suppose $u \in H^1(\Omega)$ is solution of problem (3.10). We have to show that $u \in H^1(\Omega)$ is a minimizer of the functional J_ε , which means

$$J_\varepsilon(u) \leq J_\varepsilon(v) \quad \forall v \in H^1(\Omega)$$

Therefore, let $v \in H^1(\Omega)$ be an arbitrary function. We have

$$\begin{aligned} J_\varepsilon(v) - J_\varepsilon(u) &= \frac{1}{2}a_\varepsilon(v, v) - l(v) - \frac{1}{2}a_\varepsilon(u, u) + l(u) \\ &= \frac{1}{2}a_\varepsilon(v, v) - a_\varepsilon(u, v) - \frac{1}{2}a_\varepsilon(u, u) + a_\varepsilon(u, u) \\ &= \frac{1}{2}[a_\varepsilon(u, u) - 2a_\varepsilon(u, v) + a_\varepsilon(v, v)] = \frac{1}{2}a_\varepsilon(u - v, u - v) \geq 0 \end{aligned}$$

(3.7) \Rightarrow (3.10)

Suppose $u \in H^1(\Omega)$ is a solution of problem (3.7). Then for an arbitrary $\tau \in \mathbb{R}$ and $v \in H^1(\Omega)$ we have

$$J_\varepsilon(u) \leq J_\varepsilon(u + \tau v)$$

Let us introduce a function $j : \mathbb{R} \rightarrow \mathbb{R}$ defined by $j(\tau) = J_\varepsilon(u + \tau v)$. Since j has a minimum at $\tau = 0$, then $j'(0) = 0$. We have

$$\begin{aligned} j'(0) &= \frac{d}{d\tau} \left[\frac{1}{2}a_\varepsilon(u + \tau v, u + \tau v) - l(u + \tau v) + \frac{1}{2} \int_\Omega f^2 dx \right] \Big|_{\tau=0} \\ &= \frac{d}{d\tau} \left[\frac{1}{2}a_\varepsilon(u, u) - l(u) + \tau a_\varepsilon(u, v) - \tau l(u) + \frac{\tau^2}{2}a_\varepsilon(v, v) + \frac{1}{2} \int_\Omega f^2 dx \right] \Big|_{\tau=0} \\ &= [a_\varepsilon(u, v) - l(u) + \tau a_\varepsilon(v, v)] \Big|_{\tau=0} = a_\varepsilon(u, v) - l(u) = 0 \end{aligned}$$

□

Theorem 3.2 *The problem (3.10) has a unique solution $u_\varepsilon \in H^1(\Omega)$ which depends continuously on the data.*

Proof.

The proof is a simple application of the Lax-Milgram theorem. Using the Cauchy-Schwarz inequality we obtain

$$|a_\varepsilon(u, v)| \leq |a_0(u, v)| = \langle u, v \rangle_{H^1(\Omega)} \leq \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} \quad (3.16)$$

Since $0 < \alpha < 1$ we have

$$a_\varepsilon(u, u) \geq \alpha a_0(u, u) = \alpha \|u\|_{H^1(\Omega)}^2 \quad (3.17)$$

Using the Hölder inequality and the Sobolev imbedding theorem we obtain

$$\left| \int_{\Omega} f v \, dx \right| \leq \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \leq \|f\|_{L^2(\Omega)} \|v\|_{H^1(\Omega)} \quad (3.18)$$

It is not difficult to prove that the bilinear form a_ε is symmetric. From inequalities (3.16) and (3.17), it follows that a_ε is continuous and elliptic. Continuity of the linear form l follows from inequality (3.18)

□

Corollary 3.1 *The function $u_\varepsilon \in H^1(\Omega)$ is the unique solution of problem (3.7) and problem (3.8).*

Proof.

This corollary is an immediate consequence of Theorem 3.1 and Theorem 3.2

□

Remark 3.1 *In particular, for $\varepsilon = 0$, to prove Theorem 3.2 is enough to apply the Riesz representation theorem.*

Proof.

Note that the bilinear form a_ε is nothing but the $H^1(\Omega)$ scalar product and the linear form l is the scalar product in $L^2(\Omega)$. Using that $L^\infty(\Omega) \subset L^2(\Omega) \subset$

$H^1(\Omega)'$ we can extend problem (3.10) by including $f \in H^1(\Omega)'$, then the corresponding weak formulation is given by

$$\begin{cases} \text{Find } u \in H^1(\Omega) \\ \langle u, v \rangle_{H^1(\Omega)} = \langle f, v \rangle_{H^1(\Omega)'} \quad \forall v \in H^1(\Omega) \end{cases} \quad (3.19)$$

The unique solution $u \in H^1(\Omega)$ of problem (3.19) follows immediately from the Riesz representation theorem

□

Theoretically, without lost of generality, we will further consider the case of a single inhomogeneity $B_\varepsilon = \varepsilon B$, centred in the origin $0 \in B$.

3.2 Asymptotic analysis

The purpose of this section is to prove that asymptotic expansion (3.4) holds and to derive an explicit formula for its dominant term g .

3.2.1 Preliminaries

Let us first introduce two basic definitions, commonly used in the asymptotic analysis

Definition 3.1 *We write*

$$f_1(x) = O(f_2(x)) \quad \text{as } x \rightarrow a$$

if and only if there exist a number $\delta > 0$ and constant $C > 0$, such that

$$|f_1(x)| \leq C|f_2(x)| \quad \text{for } |x - a| < \delta$$

Definition 3.2 *We write*

$$f_1(x) = o(f_2(x)) \quad \text{as } x \rightarrow a$$

if and only if

$$\lim_{x \rightarrow a} \frac{|f_1(x)|}{|f_2(x)|} = 0$$

In the first step, we derive energy estimates for difference $u_\varepsilon - u_0$. We have the following lemmas

Lemma 3.3 *For sufficiently small $\varepsilon \geq 0$, the following estimation holds*

$$\|u_\varepsilon - u_0\|_{H^1(\Omega)} = O(\varepsilon)$$

Proof.

Since the diffusivity α_ε is positive, uniformly bounded and

$$\frac{1}{\alpha} \alpha_\varepsilon = \begin{cases} 1 & \text{if } x \in B_\varepsilon \\ \frac{1}{\alpha} > 1 & \text{if } x \in \Omega \setminus \bar{B}_\varepsilon \end{cases}$$

we have

$$\int_{\Omega} |\nabla(u_\varepsilon - u_0)|^2 dx \leq \frac{1}{\alpha} \int_{\Omega} \alpha_\varepsilon |\nabla(u_\varepsilon - u_0)|^2 dx$$

$$\int_{\Omega} (u_\varepsilon - u_0)^2 dx \leq \frac{1}{\alpha} \int_{\Omega} (u_\varepsilon - u_0)^2 dx$$

From the two above inequalities and definition of the bilinear form a_ε we obtain

$$\|u_\varepsilon - u_0\|_{H^1(\Omega)}^2 \leq \frac{1}{\alpha} a_\varepsilon(u_\varepsilon - u_0, u_\varepsilon - u_0) \quad (3.20)$$

From (3.10) and properties of a bilinear form it follows that

$$\begin{aligned} a_\varepsilon(u_\varepsilon - u_0, u_\varepsilon - u_0) &= l(u_\varepsilon - u_0) - a_\varepsilon(u_0, u_\varepsilon - u_0) \\ &= a_0(u_0, u_\varepsilon - u_0) - a_\varepsilon(u_0, u_\varepsilon - u_0) \\ &= (1 - \alpha) \int_{B_\varepsilon} \nabla u_0 \cdot \nabla(u_\varepsilon - u_0) dx \end{aligned} \quad (3.21)$$

Therefore, combining (3.20) and (3.21), we obtain

$$\|u_\varepsilon - u_0\|_{H^1(\Omega)}^2 \leq \frac{1}{\alpha} (1 - \alpha) \left| \int_{B_\varepsilon} \nabla u_0 \cdot \nabla(u_\varepsilon - u_0) dx \right|$$

Application of the Cauchy-Schwarz inequality to the above integral yields

$$\begin{aligned}
\|u_\varepsilon - u_0\|_{H^1(\Omega)}^2 &\leq C \|\nabla u_0\|_{L^2(B_\varepsilon)} \|\nabla(u_\varepsilon - u_0)\|_{L^2(B_\varepsilon)} \\
&\leq C \|\nabla u_0\|_{L^2(B_\varepsilon)} \|\nabla(u_\varepsilon - u_0)\|_{L^2(\Omega)} \\
&\leq C \|\nabla u_0\|_{L^2(B_\varepsilon)} \|u_\varepsilon - u_0\|_{H^1(\Omega)}
\end{aligned}$$

Dividing both sides of inequality (3.22) by $\|u_\varepsilon - u_0\|_{H^1(\Omega)}$, we get

$$\|u_\varepsilon - u_0\|_{H^1(\Omega)} \leq C \|\nabla u_0\|_{L^2(B_\varepsilon)}$$

Finally, since ∇u_0 is uniformly bounded on B_ε , we obtain

$$\|u_\varepsilon - u_0\|_{H^1(\Omega)} \leq C\varepsilon$$

□

Lemma 3.4 *For sufficiently small $\varepsilon \geq 0$, the following estimation holds*

$$\|u_\varepsilon - u_0\|_{L^2(B_\varepsilon)} = O(\varepsilon) \quad (3.22)$$

Proof.

To prove this lemma, it is enough to note that

$$\|u_\varepsilon - u_0\|_{L^2(B_\varepsilon)} \leq \|u_\varepsilon - u_0\|_{L^2(\Omega)}$$

and then to apply the Sobolev embedding theorem and Lemma 3.3

□

3.2.2 Integral representation formula

In order to derive the integral representation of functions u_ε and u_0 , we need to know the fundamental solution of the equation $u - \Delta u = 0$. From [53, page 490] we have

Definition 3.3 *The function $\Gamma : \mathbb{R}^2 \setminus \{y\} \rightarrow \mathbb{R}$ defined by*

$$\Gamma(x, y) := \frac{1}{2\pi} K_0(|x - y|) \quad \forall x, y \in \mathbb{R}^2, \quad x \neq y \quad (3.23)$$

where K_0 denote modified Bessel function of the second kind, it is called the fundamental solution or the free space Green function of the equation

$$u - \Delta u = 0$$

In [60, page 499] one can find the sum formula for the function K_0 , given by

$$K_0(z) = -\left(\ln\left(\frac{z}{2}\right) + \gamma\right) I_0(z) + \frac{\frac{1}{4}z^2}{(1!)^2} + \left(1 + \frac{1}{2}\right) \frac{(\frac{1}{4}z^2)^2}{(2!)^2} + \dots \quad (3.24)$$

where

$$I_0(z) = 1 + \frac{\frac{1}{4}z^2}{(1!)^2} + \left(1 + \frac{1}{2}\right) \frac{(\frac{1}{4}z^2)^2}{(2!)^2} + \left(1 + \frac{1}{2} + \frac{1}{3}\right) \frac{(\frac{1}{4}z^2)^3}{(3!)^2} + \dots \quad (3.25)$$

is the modified Bessel function of the first kind and γ denote the Euler-Mascheroni constant.

From expansions (3.24) and (3.25) we deduce the following asymptotic

$$K_0(z) = -\ln z + \ln 2 - \gamma + O(z^2 \ln z) \quad \text{for } z \rightarrow 0 \quad (3.26)$$

Therefore, the fundamental solution $\Gamma(\cdot, y)$, defined in (3.23), has the same singular behavior as the fundamental solution

$$\Phi(x, y) = -\frac{1}{2\pi} \ln(|x - y|) \quad (3.27)$$

of the Laplace equation $\Delta u = 0$.

From (3.23), (3.26) and (3.27), for $|x - y| \rightarrow 0$, we obtain

$$\Gamma(x, y) = \Phi(x, y) + \ln 2 - \gamma + O(|x - y|^2 \ln |x - y|) \quad (3.28)$$

In order to see how to derive the integral representation formula for a solution of the Poisson equation, we refer to Evans [29, page 23]. For the case of solution u_0 of equation $u_0 - \Delta u_0 = f$, we may proceed analogously, making use of the estimation (3.28). This representation, for any $y \in \Omega$, is given by

$$\begin{aligned} u_0(y) &= \int_{\partial\Omega} \left[u_0(x) \frac{\partial \Gamma}{\partial n}(x, y) - \Gamma(x, y) \frac{\partial u_0}{\partial n}(x) \right] ds(x) \\ &\quad + \int_{\Omega} \Gamma(x, y) f(x) dx \end{aligned} \quad (3.29)$$

In a similar manner, we can derive an integral representation of the function u_ε , solution of the problem (3.8) for any $y \in \Omega \setminus \bar{B}_\varepsilon$. We have

$$\begin{aligned} u_\varepsilon(y) &= \int_{\partial\Omega} \left[u_\varepsilon(x) \frac{\partial\Gamma}{\partial n}(x, y) - \Gamma(x, y) \frac{\partial u_\varepsilon}{\partial n}(x) \right] ds(x) \\ &\quad + \int_{\partial B_\varepsilon} \left[u_\varepsilon(x) \frac{\partial\Gamma}{\partial n}(x, y) - \Gamma(x, y) \frac{\partial u_\varepsilon^+}{\partial n}(x) \right] ds(x) \\ &\quad + \int_{\Omega \setminus \bar{B}_\varepsilon} \Gamma(x, y) f(x) dx \end{aligned} \quad (3.30)$$

Subtracting (3.29) from (3.30), we get

$$\begin{aligned} u_\varepsilon(y) - u_0(y) &= \int_{\partial B_\varepsilon} \left[u_\varepsilon(x) \frac{\partial\Gamma}{\partial n}(x, y) - \Gamma(x, y) \frac{\partial u_\varepsilon^+}{\partial n}(x) \right] ds(x) \\ &\quad + \int_{\partial\Omega} [u_\varepsilon(x) - u_0(x)] \frac{\partial\Gamma}{\partial n}(x, y) ds(x) - \int_{B_\varepsilon} \Gamma(x, y) f(x) dx \end{aligned} \quad (3.31)$$

Let us now introduce the Green function $G(\cdot, y)$ for the Neumann problem

$$\begin{cases} u_0 - \Delta u_0 = f & x \in \Omega \\ \frac{\partial u_0}{\partial n} = 0 & x \in \partial\Omega \end{cases} \quad (3.32)$$

This function is the unique solution of the problem

$$\begin{cases} G(x, y) - \Delta_x G(x, y) = \delta_y(x) & x \in \Omega \\ \frac{\partial G}{\partial n}(x, y) = 0 & x \in \partial\Omega \end{cases}$$

and can be expressed as a sum of the fundamental solution $\Gamma(\cdot, y)$ and the so called corrector function $h(\cdot, y)$, which is chosen in such a way that for $x \in \partial\Omega$ and $y \in \Omega$, the following is true

$$\frac{\partial G}{\partial n}(x, y) = 0 \quad (3.33)$$

Thus, we can write

$$G(x, y) = \Gamma(x, y) + h(x, y) \quad (3.34)$$

where the function $h(\cdot, y)$ satisfies the following boundary value problem

$$\begin{cases} h(x, y) - \Delta_x h(x, y) = 0 & x \in \Omega \\ \frac{\partial h}{\partial n}(x, y) = -\frac{\partial \Gamma}{\partial n}(x, y) & x \in \partial\Omega \end{cases}$$

Using the Green formula, we can show that function $h(\cdot, y)$ satisfies

$$\begin{aligned} 0 &= \int_{\partial B_\varepsilon} \left[u_\varepsilon(x) \frac{\partial h}{\partial n}(x, y) - h(x, y) \frac{\partial u_\varepsilon^+}{\partial n}(x) \right] ds(x) \\ &\quad + \int_{\partial\Omega} [u_\varepsilon(x) - u_0(x)] \frac{\partial h}{\partial n}(x, y) ds(x) - \int_{B_\varepsilon} h(x, y) f(x) dx \end{aligned} \quad (3.35)$$

Summing (3.31) and (3.35) both sides and using (3.34) and (3.33), we get

$$\begin{aligned} u_\varepsilon(y) - u_0(y) &= \int_{\partial B_\varepsilon} \left[u_\varepsilon(x) \frac{\partial G}{\partial n}(x, y) - G(x, y) \frac{\partial u_\varepsilon^+}{\partial n}(x) \right] ds(x) \\ &\quad - \int_{B_\varepsilon} G(x, y) f(x) dx \end{aligned} \quad (3.36)$$

for any $y \in \Omega \setminus \bar{B}_\varepsilon$.

3.2.3 Estimation of the Green function

We will need the following estimation

Lemma 3.5 *For sufficiently small $\varepsilon \geq 0$ and $y \in B_{2\varepsilon} \setminus \bar{B}_\varepsilon$ the following estimation holds*

$$\|G(\cdot, y)\|_{L^2(B_\varepsilon)} = O(\varepsilon |\ln \varepsilon|)$$

Proof.

From the Minkowski inequality, equation (3.34) and estimation (3.28), we obtain

$$\begin{aligned} \|G(\cdot, y)\|_{L^2(B_\varepsilon)} &= \|\Gamma(\cdot, y) + h(\cdot, y)\|_{L^2(B_\varepsilon)} \leq \|\Phi(\cdot, y)\|_{L^2(B_\varepsilon)} + C\varepsilon \\ &\quad + C\|\xi(\cdot, y)\|_{L^2(B_\varepsilon)} + \|h(\cdot, y)\|_{L^2(B_\varepsilon)} \end{aligned}$$

where $\xi(x, y) := -|x - y|^2 \ln|x - y|$ for sufficiently small $|x - y| < 1$.

First, let us find an upper bound of the norm

$$\|\Phi(\cdot, y)\|_{L^2(B_\varepsilon)} = \frac{1}{2\pi} \sqrt{\int_{B_\varepsilon} \ln^2|x - y| dx}$$

where B_ε and $B_{2\varepsilon}$ are open balls of radius ε and 2ε , respectively, centred in the origin $0 \in B_\varepsilon$. The point $y \in B_{2\varepsilon} \setminus \bar{B}_\varepsilon$ is fixed.

Let $x = (x_1, x_2)$ and $y = (y_1, y_2)$. Applying a change of variables

$$\begin{cases} x_1 = r \cos \theta & 0 \leq r < \varepsilon \\ x_2 = r \sin \theta & 0 \leq \theta \leq 2\pi \end{cases} \quad \begin{cases} y_1 = R \cos \beta & \varepsilon < R < 2\varepsilon \\ y_2 = R \sin \beta & 0 \leq \beta \leq 2\pi \end{cases} \quad (3.37)$$

we obtain

$$|x - y| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} = \sqrt{r^2 + R^2 - 2rR \cos(\theta - \beta)}$$

We have

$$\int_{B_\varepsilon} \ln^2|x - y| dx = \int_0^{2\pi} \int_0^\varepsilon r \ln^2 \sqrt{r^2 + R^2 - 2rR \cos(\theta - \beta)} dr d\theta$$

Since

$$0 < R - r \leq \sqrt{r^2 + R^2 - 2rR \cos(\theta - \beta)} \leq R + r < 3\varepsilon$$

and $\ln^2(\cdot)$ is a strictly decreasing function on the interval $[0, 1]$, for $\varepsilon \leq 3^{-1}$, we obtain

$$\int_{B_\varepsilon} \ln^2|x - y| dx \leq \int_0^{2\pi} \int_0^\varepsilon r \ln^2(R - r) dr d\theta$$

We have

$$\int_0^\varepsilon r \ln^2(R - r) dr < \lim_{a \rightarrow \varepsilon^-} \int_0^a r \ln^2(\varepsilon - r) dr = \frac{7}{4}\varepsilon^2 - \frac{3}{2}\varepsilon^2 \ln \varepsilon + \frac{1}{2}\varepsilon^2 \ln^2 \varepsilon$$

Therefore, for $0 < \varepsilon \leq e^{-3}$, we have

$$\frac{7}{4}\varepsilon^2 - \frac{3}{2}\varepsilon^2 \ln \varepsilon + \frac{1}{2}\varepsilon^2 \ln^2 \varepsilon \leq C\varepsilon^2 \ln^2 \varepsilon$$

Finally, we obtain

$$\|\Phi(x, y)\|_{L^2(B_\varepsilon)} \leq C \sqrt{\int_{B_\varepsilon} \ln^2 |x - y| dx} \leq C\varepsilon |\ln \varepsilon|$$

In the similar manner as above, one can show that

$$\|\xi(\cdot, y)\|_{L^2(B_\varepsilon)} \leq C\varepsilon^3 |\ln \varepsilon|$$

Since function $\Gamma(\cdot, y)$ is in $C^\infty(\Omega)$ for $x \neq y$, we have $\|h(\cdot, y)\|_{L^2(B_\varepsilon)} = O(\varepsilon)$

□

In the proofs of the two next lemmas we follow Vogelius et al. [61].

Lemma 3.6 *For sufficiently small $\varepsilon \geq 0$ and $y \in B_{2\varepsilon} \setminus \bar{B}_\varepsilon$, we have the following estimation*

$$u_\varepsilon(y) - u_0(y) = (1 - \alpha) \int_{\partial B_\varepsilon} u_\varepsilon(x) \frac{\partial G}{\partial n}(x, y) ds(x) + O(\varepsilon^2 |\ln \varepsilon|)$$

Proof.

Applying the condition (3.9), the divergence theorem and the Green formula, we obtain

$$\begin{aligned} \int_{\partial B_\varepsilon} G(x, y) \frac{\partial u_\varepsilon^+}{\partial n}(x) ds(x) &= \alpha \int_{\partial B_\varepsilon} G(x, y) \frac{\partial u_\varepsilon^-}{\partial n}(x) ds(x) \\ &= \alpha \int_{B_\varepsilon} \nabla \cdot [G(x, y) \nabla u_\varepsilon(x)] dx \\ &= \alpha \int_{B_\varepsilon} G(x, y) \Delta u_\varepsilon(x) dx + \alpha \int_{B_\varepsilon} \nabla G(x, y) \cdot \nabla u_\varepsilon(x) dx \\ &= \int_{B_\varepsilon} G(x, y) [u_\varepsilon(x) - f(x)] dx - \alpha \int_{B_\varepsilon} \Delta G(x, y) u_\varepsilon(x) dx \\ &\quad + \alpha \int_{\partial B_\varepsilon} u_\varepsilon(x) \frac{\partial G}{\partial n}(x, y) ds(x) \end{aligned}$$

Since $y \in B_{2\varepsilon} \setminus \bar{B}_\varepsilon$, we have $G(x, y) - \Delta G(x, y) = 0$, for any $x \in B_\varepsilon$. Combining this fact with the above equality, we obtain

$$\begin{aligned} \int_{\partial B_\varepsilon} G(x, y) \frac{\partial u_\varepsilon^+}{\partial n}(x) ds(x) &= (1 - \alpha) \int_{B_\varepsilon} G(x, y) u_\varepsilon(x) dx \\ &\quad - \int_{B_\varepsilon} G(x, y) f(x) dx + \alpha \int_{\partial B_\varepsilon} u_\varepsilon(x) \frac{\partial G}{\partial n}(x, y) ds(x) \end{aligned} \quad (3.38)$$

Substituting (3.38) into (3.36), we get

$$\begin{aligned} u_\varepsilon(y) - u_0(y) &= (1 - \alpha) \int_{\partial B_\varepsilon} u_\varepsilon(x) \frac{\partial G}{\partial n}(x, y) ds(x) \\ &\quad - (1 - \alpha) \int_{B_\varepsilon} u_\varepsilon(x) G(x, y) dx \end{aligned} \quad (3.39)$$

To complete the proof, what remains to be checked is

$$\int_{B_\varepsilon} u_\varepsilon(x) G(x, y) dx \leq C\varepsilon^2 |\ln \varepsilon|$$

From the Minkowski inequality and Lemma 3.4 we obtain

$$\|u_\varepsilon\|_{L^2(B_\varepsilon)} \leq \|u_\varepsilon - u_0\|_{L^2(B_\varepsilon)} + \|u_0\|_{L^2(B_\varepsilon)} \leq C\varepsilon \quad (3.40)$$

The Hölder inequality and Lemma 3.5 yields

$$\int_{B_\varepsilon} u_\varepsilon(x) G(x, y) dx \leq \|u_\varepsilon\|_{L^2(B_\varepsilon)} \|G(\cdot, y)\|_{L^2(B_\varepsilon)} \leq C\varepsilon^2 |\ln \varepsilon|$$

□

Lemma 3.7 *For sufficiently small $\varepsilon \geq 0$ and $y \in B_{2\varepsilon} \setminus \bar{B}_\varepsilon$, we have the following estimation*

$$\int_{\partial B_\varepsilon} u_\varepsilon(x) \frac{\partial G}{\partial n}(x, y) ds(x) = \int_{\partial B_\varepsilon} u_\varepsilon(x) \frac{\partial \Phi}{\partial n}(x, y) ds(x) + O(\varepsilon^2 |\ln \varepsilon|)$$

Proof.

From (3.34) and next (3.28), we have

$$\int_{\partial B_\varepsilon} u_\varepsilon(x) \frac{\partial G}{\partial n}(x, y) ds(x) = \int_{\partial B_\varepsilon} u_\varepsilon(x) \frac{\partial \Gamma}{\partial n}(x, y) ds(x) + \int_{\partial B_\varepsilon} u_\varepsilon(x) \frac{\partial h}{\partial n}(x, y) ds(x)$$

$$\int_{\partial B_\varepsilon} u_\varepsilon(x) \frac{\partial \Gamma}{\partial n}(x, y) ds(x) \leq \int_{\partial B_\varepsilon} u_\varepsilon(x) \frac{\partial \Phi}{\partial n}(x, y) ds(x) + C \int_{\partial B_\varepsilon} u_\varepsilon(x) \frac{\partial \xi}{\partial n}(x, y) ds(x)$$

where function $\xi(\cdot, y)$ is defined by $\xi(x, y) := -|x - y|^2 \ln|x - y|$ for sufficiently small $|x - y| < 1$.

Thus, we have to find estimations of the two last integrals of the above expressions.

Since the function $G(\cdot, y)$ satisfies $G(x, y) - \Delta_x G(x, y) = 0$ for $x \in B_\varepsilon$ and $y \in B_{2\varepsilon} \setminus \bar{B}_\varepsilon$, we have

$$G(x, y) = \Delta_x h(x, y) + \Delta_x \Gamma(x, y) \quad (3.41)$$

Using the divergence theorem and equality (3.41), we obtain

$$\int_{\partial B_\varepsilon} \frac{\partial h}{\partial n}(x, y) ds(x) = \int_{B_\varepsilon} \Delta_x h(x, y) dx = \int_{B_\varepsilon} G(x, y) dx - \int_{B_\varepsilon} \Delta_x \Gamma(x, y) dx \quad (3.42)$$

Since function $\Phi(\cdot, y)$ satisfies $\Delta_x \Phi(x, y) = 0$ for $x \in B_\varepsilon$ and $y \in B_{2\varepsilon} \setminus \bar{B}_\varepsilon$ and function $-\Delta_x \Gamma(\cdot, y)$ is bounded from above by 0, using Lemma 3.5, we obtain

$$\int_{\partial B_\varepsilon} \frac{\partial h}{\partial n}(x, y) ds(x) \leq \int_{B_\varepsilon} G(x, y) dx \leq C\varepsilon \|G(\cdot, y)\|_{L^2(B_\varepsilon)} \leq C\varepsilon^2 |\ln \varepsilon| \quad (3.43)$$

Using this fact, we can write

$$\int_{\partial B_\varepsilon} u_\varepsilon(x) \frac{\partial h}{\partial n}(x, y) ds(x) = \int_{\partial B_\varepsilon} [u_\varepsilon(x) - u_0(y)] \frac{\partial h}{\partial n}(x, y) ds(x) + O(\varepsilon^2 |\ln \varepsilon|)$$

and next

$$\begin{aligned} \int_{\partial B_\varepsilon} u_\varepsilon(x) \frac{\partial h}{\partial n}(x, y) ds(x) &= \int_{\partial B_\varepsilon} [u_\varepsilon(x) - u_0(x)] \frac{\partial h}{\partial n}(x, y) ds(x) \\ &\quad + \int_{\partial B_\varepsilon} [u_0(x) - u_0(y)] \frac{\partial h}{\partial n}(x, y) ds(x) + O(\varepsilon^2 |\ln \varepsilon|) \end{aligned} \quad (3.44)$$

Using the divergence theorem to the first integral on the right hand side of the above equation, we get

$$\begin{aligned} \int_{\partial B_\varepsilon} [u_\varepsilon(x) - u_0(x)] \frac{\partial h}{\partial n}(x, y) ds(x) &= \int_{B_\varepsilon} [u_\varepsilon(x) - u_0(x)] \Delta h(x, y) dx \\ &\quad + \int_{B_\varepsilon} \nabla [u_\varepsilon(x) - u_0(x)] \cdot \nabla h(x, y) dx \end{aligned}$$

Next, application of (3.41) and the Hölder inequality yields

$$\begin{aligned} \int_{\partial B_\varepsilon} [u_\varepsilon(x) - u_0(x)] \frac{\partial h}{\partial n}(x, y) ds(x) &\leq \|u_\varepsilon - u_0\|_{L^2(B_\varepsilon)} \|G(\cdot, y)\|_{L^2(B_\varepsilon)} \\ &\quad + \|\nabla(u_\varepsilon - u_0)\|_{L^2(B_\varepsilon)} \|\nabla h(\cdot, y)\|_{L^2(B_\varepsilon)} \end{aligned} \quad (3.45)$$

Using Lemma 3.4, Lemma 3.3 and Lemma 3.5, we obtain

$$\int_{\partial B_\varepsilon} [u_\varepsilon(x) - u_0(x)] \frac{\partial h}{\partial n}(x, y) ds(x) \leq C\varepsilon^2 |\ln \varepsilon| + C\varepsilon^2 \leq C\varepsilon^2 |\ln \varepsilon| \quad (3.46)$$

In order to bound the second integral on the right hand side of equation (3.44), we use estimation (3.43). Thus we have

$$\begin{aligned} \int_{\partial B_\varepsilon} [u_0(x) - u_0(y)] \frac{\partial h}{\partial n}(x, y) ds(x) \\ \leq \|u_0(\cdot) - u_0(y)\|_{L^\infty(\partial B_\varepsilon)} \left\| \frac{\partial h}{\partial n}(\cdot, y) \right\|_{L^1(\partial B_\varepsilon)} \leq C\varepsilon^2 |\ln \varepsilon| \end{aligned} \quad (3.47)$$

In order to estimate integral

$$\int_{\partial B_\varepsilon} u_\varepsilon(x) \frac{\partial \Gamma}{\partial n}(x, y) ds(x) \quad (3.48)$$

we need to find an estimate of the following expression

$$\int_{\partial B_\varepsilon} \frac{\partial \xi}{\partial n}(x, y) ds(x) = \int_{B_\varepsilon} \Delta_x \xi(x, y) dx \quad (3.49)$$

and then to proceed analogously as in estimation of the integral (3.44).

Applying changing of variables as in (3.37) and calculate laplacian of the function $\xi(\cdot, y)$ in polar coordinates, we obtain

$$\Delta \xi(r, \theta) = -4 - 2 \ln[r^2 + R^2 - 2rR \cos(\theta - \beta)]$$

Using that

$$-4 - 2 \ln[r^2 + R^2 - 2rR \cos(\theta - \beta)] \leq -4 - 2 \ln(R - r)$$

we get

$$\int_0^\varepsilon r \Delta \xi(r, \theta) dr < \lim_{a \rightarrow \varepsilon^-} \int_0^a r [-4 - 2 \ln(\varepsilon - r)] dr = -\frac{1}{2} \varepsilon^2 - \varepsilon^2 \ln \varepsilon$$

Thus, for sufficiently small ε , we have

$$\int_{\partial B_\varepsilon} \frac{\partial \xi}{\partial n}(x, y) ds(x) = O(\varepsilon^2 |\ln \varepsilon|) \quad (3.50)$$

□

From Lemma 3.6 and Lemma 3.10, we get

$$u_\varepsilon(y) - u_0(y) = (1 - \alpha) \int_{\partial B_\varepsilon} u_\varepsilon(x) \frac{\partial \Phi}{\partial n}(x, y) ds(x) + O(\varepsilon^2 |\ln \varepsilon|) \quad (3.51)$$

3.2.4 Asymptotic expansion of $u_\varepsilon - u_0$ on the boundary of the inhomogeneity

Based on estimates of H^1 -norm and L^2 -norm of expression $u_\varepsilon - u_0$, and boundary integral formulation (3.51), we are now able to obtain an asymp-

otic formula for $u_\varepsilon - u_0$ on the boundary of the inhomogeneity B_ε .

Let us first introduce the definition of the single and double layer potentials [42, page 67]

Definition 3.4 *Given an integrable function φ , the integrals*

$$u(y) = \int_{\partial B} \varphi(x) \Phi(x, y) ds(x), \quad y \in \mathbb{R}^N \setminus \partial B$$

and

$$v(y) = \int_{\partial B} \varphi(x) \frac{\partial \Phi}{\partial n}(x, y) ds(x), \quad y \in \mathbb{R}^N \setminus \partial B$$

are called, single layer and double layer potentials, respectively, with density φ .

For the investigation of the boundary value problems we need properties of these potentials for points on the boundary of inhomogeneity B_ε where the integrals become singular. These properties are given in the following theorem

Theorem 3.3 *Let ∂B be of class C^2 and let $\varphi \in C(\partial B)$. Then the single layer potential u with density φ is continuous throughout \mathbb{R}^N . On the boundary, it holds*

$$u(y) = \int_{\partial B} \varphi(x) \Phi(x, y) ds(x), \quad y \in \partial B$$

and

$$\frac{\partial u^\pm}{\partial n}(y) = \int_{\partial B} \varphi(x) \frac{\partial \Phi}{\partial n}(x, y) ds(x) \mp \frac{1}{2} \varphi(y) \quad y \in \partial B$$

where

$$\frac{\partial u^\pm}{\partial n}(y) = \lim_{\rho \rightarrow 0} [\nabla u(y \pm \rho n(y)) \cdot n(y)]$$

is to be understood in a sense of uniform convergence on ∂B and where the integrals exist as improper integrals. The double layer potential v with density φ can be continuously extended from B^+ to \bar{B}^+ and from B^- to \bar{B}^- with limiting values

$$v^\pm(y) = \int_{\partial B} \varphi(x) \frac{\partial \Phi}{\partial n}(x, y) ds(x) \pm \frac{1}{2} \varphi(y) \quad y \in \partial B \quad (3.52)$$

where

$$v^\pm(y) = \lim_{\rho \rightarrow 0} v(y \pm \rho n(y))$$

and where the integrals exist as improper integrals. Furthermore

$$\lim_{\rho \rightarrow 0} \left[\frac{\partial v}{\partial n}(y + \rho n(y)) - \frac{\partial v}{\partial n}(y - \rho n(y)) \right] = 0 \quad y \in \partial B$$

uniformly on ∂B

Proof.

The proof of these so-called jump relations for the single and double layer potentials is very elaborate and lengthy. One can find the scratch of this proof in Colton and Kress [25, page 40]. Refer to Colton [24] if interested in details.

□

The immediate consequence of Theorem 3.3 is the following corollary

Corollary 3.2 *For the double layer potential with constant density there holds*

$$2 \int_{\partial B} \frac{\partial \Phi}{\partial n}(x, y) ds(x) = \begin{cases} -2 & y \in B^- \\ -1 & y \in \partial B \\ 0 & y \in B^+ \end{cases} \quad (3.53)$$

Proof.

One can find the proof in [42, page 68]

□

Let us introduce the auxiliary, vector valued function $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, which is a solution of the problem

$$\begin{cases} \Delta\phi = 0 & x \in \mathbb{R}^2 \setminus \bar{B} \text{ and } x \in B \\ \phi^+ = \phi^- & x \in \partial B \\ \frac{1}{\alpha} \frac{\partial\phi^+}{\partial n} - \frac{\partial\phi^-}{\partial n} = -n & x \in \partial B \\ \lim_{x \rightarrow \infty} \phi(x) = 0 \end{cases} \quad (3.54)$$

The existence and uniqueness of its solution, and bounds of the function ϕ as $x \rightarrow \infty$, given by

$$\phi(x) = O(|x|^{-1}) \quad \text{and} \quad \nabla\phi(x) = O(|x|^{-2}) \quad (3.55)$$

are established in [21]. Following authors, we will derive the integral representation of the function ϕ for $\tilde{y} \in \partial B$. We have the following lemma

Lemma 3.8 *For $\tilde{y} \in \partial B$, the function ϕ has the following integral representation*

$$\frac{1}{2}(1 + \alpha)\phi(\tilde{y}) = (1 - \alpha) \int_{\partial B} \phi(\tilde{x}) \frac{\partial\Phi}{\partial n}(\tilde{x}, \tilde{y}) ds(\tilde{x}) + \alpha \int_{\partial B} n \Phi(\tilde{x}, \tilde{y}) ds(\tilde{x})$$

Proof.

Let $x \in \mathbb{R}^2$ and let B_R be a disk of radius R , such that $B \subset B_R$ and $x \in B_R$. Applying the Green formula for the domain $B_R \setminus \bar{B}$, we obtain

$$\begin{aligned} \phi(\tilde{y}) &= \int_{\partial B_R} \frac{\partial\phi}{\partial n}(\tilde{x}) \Phi(\tilde{x}, \tilde{y}) ds(\tilde{x}) - \int_{\partial B} \frac{\partial\phi^+}{\partial n}(\tilde{x}) \Phi(\tilde{x}, \tilde{y}) ds(\tilde{x}) \\ &\quad - \int_{\partial B_R} \phi(\tilde{x}) \frac{\partial\Phi}{\partial n}(\tilde{x}, \tilde{y}) ds(\tilde{x}) + \int_{\partial B} \phi(\tilde{x}) \frac{\partial\Phi}{\partial n}(\tilde{x}, \tilde{y}) ds(\tilde{x}) \end{aligned} \quad (3.56)$$

From (3.55), we have

$$\phi(y) = \int_{\partial B} \phi(\tilde{x}) \frac{\partial\Phi}{\partial n}(\tilde{x}, \tilde{y}) ds(\tilde{x}) - \int_{\partial B} \frac{\partial\phi^+}{\partial n}(\tilde{x}) \Phi(\tilde{x}, \tilde{y}) ds(\tilde{x})$$

Using the boundary condition

$$\frac{1}{\alpha} \frac{\partial\phi^+}{\partial n}(\tilde{x}) - \frac{\partial\phi^-}{\partial n}(\tilde{x}) = -n(\tilde{x}) \quad \text{for } \tilde{x} \in \partial B$$

we obtain

$$\begin{aligned} \phi(\tilde{y}) &= \int_{\partial B} \phi(\tilde{x}) \frac{\partial \Phi}{\partial n}(\tilde{x}, \tilde{y}) ds(\tilde{x}) + \alpha \int_{\partial B} n \Phi(\tilde{x}, \tilde{y}) ds(\tilde{x}) \\ &\quad - \alpha \int_{\partial B} \frac{\partial \phi^-}{\partial n}(\tilde{x}) \Phi(\tilde{x}, \tilde{y}) ds(\tilde{x}) \end{aligned} \quad (3.57)$$

From the Green formula applied to functions ϕ and $\Phi(\cdot, \tilde{y})$ for the domain B , we obtain

$$\int_{\partial B} \frac{\partial \phi^-}{\partial n}(\tilde{x}) \Phi(\tilde{x}, \tilde{y}) ds(\tilde{x}) = \int_{\partial B} \phi(\tilde{x}) \frac{\partial \Phi}{\partial n}(\tilde{x}, \tilde{y}) ds(\tilde{x}) \quad (3.58)$$

Substitution (3.58) into (3.57) yields

$$\phi(\tilde{y}) = (1 - \alpha) \int_{\partial B} \phi(\tilde{x}) \frac{\partial \Phi}{\partial n}(\tilde{x}, \tilde{y}) ds(\tilde{x}) + \alpha \int_{\partial B} n \Phi(\tilde{x}, \tilde{y}) ds(\tilde{x})$$

According to Definition 3.4, we know that ϕ is a double layer potential with density $(1 - \alpha)\phi$. Using Theorem 3.3 for $\tilde{y} \in \partial B$ we obtain

$$\phi(y) = (1 - \alpha) \int_{\partial B} \phi(\tilde{x}) \frac{\partial \Phi}{\partial n}(\tilde{x}, \tilde{y}) ds(\tilde{x}) + \frac{1}{2}(1 - \alpha)\phi(\tilde{y}) + \alpha \int_{\partial B} n \Phi(\tilde{x}, \tilde{y}) ds(\tilde{x})$$

Therefore, for $y \in \partial B$, we have

$$\frac{1}{2}(1 + \alpha)\phi(\tilde{y}) = (1 - \alpha) \int_{\partial B} \phi(\tilde{x}) \frac{\partial \Phi}{\partial n}(\tilde{x}, \tilde{y}) ds(\tilde{x}) + \alpha \int_{\partial B} n \Phi(\tilde{x}, \tilde{y}) ds(\tilde{x})$$

which gives the desired result

□

From the boundary integral formula (3.51), we know that the function $u_\varepsilon - u_0$ is a double layer potential with density $(1 - \alpha)u_\varepsilon$. Thus, using Theorem 3.3 and Corollary 3.2 for $y \in \partial B_\varepsilon$, we obtain

$$\begin{aligned}
& \frac{1}{2}(1 + \alpha)[u_\varepsilon(y) - u_0(y)] \\
&= (1 - \alpha) \int_{\partial B_\varepsilon} [u_\varepsilon(x) - u_0(y)] \frac{\partial \Phi}{\partial n}(x, y) ds(x) + O(\varepsilon^2 |\ln \varepsilon|)
\end{aligned} \tag{3.59}$$

Given the approximation of the Green function $G(\cdot, y)$ by the fundamental solution $\Phi(\cdot, y)$, which is defined for all $x \in \mathbb{R}^2 \setminus \{y\}$, we can directly apply the result of Vogelius et al. presented in [61]. Authors, starting with expression (3.59) prove the following theorem

Theorem 3.4 *For sufficiently small $\varepsilon \geq 0$ and $\tilde{y} \in \partial B$, we have*

$$u_\varepsilon(\varepsilon \tilde{y}) - u_0(\varepsilon \tilde{y}) = \varepsilon \frac{1}{\alpha} (1 - \alpha) \phi(\tilde{y}) \cdot \nabla u_0(0) + O(\varepsilon^2 |\ln \varepsilon|)$$

Proof.

Here we present the main idea of the proof. For the complete proof, refer to [61]. We have

$$\begin{aligned}
\frac{1}{2}(1 + \alpha)[u_\varepsilon(y) - u_0(y)] &= (1 - \alpha) \int_{\partial B_\varepsilon} [u_\varepsilon(x) - u_0(x)] \frac{\partial \Phi}{\partial n}(x, y) ds(x) \\
&+ (1 - \alpha) \int_{\partial B_\varepsilon} [u_0(x) - u_0(y)] \frac{\partial \Phi}{\partial n}(x, y) ds(x) + O(\varepsilon^2 |\ln \varepsilon|)
\end{aligned} \tag{3.60}$$

We introduce a change of variables $\tilde{x} = x/\varepsilon$, $\tilde{y} = y/\varepsilon$ and note that

$$\frac{\partial \Phi}{\partial n}(x, y) = -\frac{1}{2\pi} \frac{x - y}{|x - y|^2} = -\frac{1}{2\pi} \frac{1}{\varepsilon} \frac{\tilde{x} - \tilde{y}}{|\tilde{x} - \tilde{y}|^2} = \frac{1}{\varepsilon} \frac{\partial \Phi}{\partial n}(\tilde{x}, \tilde{y})$$

Furthermore, we have

$$|u_0(\varepsilon \tilde{x}) - u_0(\varepsilon \tilde{y}) - \varepsilon \nabla u_0(0) \cdot (\tilde{x} - \tilde{y})| = O(\varepsilon^2)$$

Hence

$$\begin{aligned}
\frac{1}{2}(1 + \alpha)[u_\varepsilon(\varepsilon \tilde{y}) - u_0(\varepsilon \tilde{y})] &= (1 - \alpha) \int_{\partial B} [u_\varepsilon(\varepsilon \tilde{x}) - u_0(\varepsilon \tilde{y})] \frac{\partial \Phi}{\partial n}(\tilde{x}, \tilde{y}) ds(\tilde{x}) \\
&+ \varepsilon (1 - \alpha) \nabla u_0(0) \cdot \int_{\partial B} (\tilde{x} - \tilde{y}) \frac{\partial \Phi}{\partial n}(\tilde{x}, \tilde{y}) ds(\tilde{x}) + O(\varepsilon^2 |\ln \varepsilon|)
\end{aligned} \tag{3.61}$$

Next, we note that

$$\begin{aligned} \int_{\partial B} (\tilde{x} - \tilde{y}) \frac{\partial \Phi}{\partial n}(\tilde{x}, \tilde{y}) ds(\tilde{x}) &= \int_B \nabla_{\tilde{x}}(\tilde{x} - \tilde{y}) \cdot \nabla_{\tilde{x}} \Phi(\tilde{x}, \tilde{y}) ds(\tilde{x}) \\ &= \int_{\partial B} n \Phi(\tilde{x}, \tilde{y}) ds(\tilde{x}) \end{aligned} \quad (3.62)$$

Substituting (3.62) into (3.61), we can write

$$\begin{aligned} \frac{1}{2}(1 + \alpha)[u_\varepsilon(\varepsilon \tilde{y}) - u_0(\varepsilon \tilde{y})] &= (1 - \alpha) \int_{\partial B} [u_\varepsilon(\varepsilon \tilde{x}) - u_0(\varepsilon \tilde{y})] \frac{\partial \Phi}{\partial n}(\tilde{x}, \tilde{y}) ds(\tilde{x}) \\ &\quad + \varepsilon(1 - \alpha) \nabla u_0(0) \int_{\partial B} n \Phi(\tilde{x}, \tilde{y}) ds(\tilde{x}) + O(\varepsilon^2 |\ln \varepsilon|) \end{aligned} \quad (3.63)$$

From Lemma 3.8, we have

$$\frac{1}{2}(1 + \alpha)\phi(\tilde{y}) = (1 - \alpha) \int_{\partial B} \phi(\tilde{x}) \frac{\partial \Phi}{\partial n}(\tilde{x}, \tilde{y}) ds(\tilde{x}) + \alpha \int_{\partial B} n \Phi(\tilde{x}, \tilde{y}) ds(\tilde{x}) \quad (3.64)$$

Next, we multiply equation (3.64) by expression $\varepsilon \frac{1}{\alpha}(1 - \alpha) \nabla u_0(0)$ and subtract the result from both sides of the equation (3.63). We obtain

$$\frac{1}{2}(1 + \alpha)\psi^*(\tilde{y}) = (1 - \alpha) \int_{\partial B} \psi^*(\tilde{x}) \frac{\partial \Phi}{\partial n}(\tilde{x}, \tilde{y}) ds(\tilde{x}) + O(\varepsilon^2 |\ln \varepsilon|) \quad (3.65)$$

where

$$\psi^*(\tilde{y}) := u_\varepsilon(\varepsilon \tilde{y}) - u_0(\varepsilon \tilde{y}) - \varepsilon \frac{1}{\alpha}(1 - \alpha) \nabla u_0(0) \cdot \phi(\tilde{y})$$

The Fredholm theory implies that the bounded linear operator

$$\psi \rightarrow (c + K)\psi$$

given by

$$(c + K)(\psi)(\tilde{y}) := \frac{1}{2}(1 + \alpha)\psi(\tilde{y}) + (1 - \alpha) \int_{\partial B} \psi(\tilde{x}) \frac{\partial \Phi}{\partial n}(\tilde{x}, \tilde{y}) ds(\tilde{x}) \quad (3.66)$$

maps $C^0(\partial B)$ onto $C^0(\partial B)$, and therefore has a bounded inverse. Due to the existence of bounded inverse for $c + K$, it now follows that

$$\begin{aligned} \|u_\varepsilon(\varepsilon \cdot) - u_0(\varepsilon \cdot) - \varepsilon \frac{1}{\alpha}(1 - \alpha) \nabla u_0(0) \cdot \phi(\cdot)\|_{C^0(\partial B)} &= \|\psi^*\|_{C^0(\partial B)} \\ &= \|(c + K)^{-1} O(\varepsilon^2 |\ln \varepsilon|)\|_{C^0(\partial B)} \leq O(\varepsilon^2 |\ln \varepsilon|) \end{aligned} \quad (3.67)$$

which gives the desired result

□

3.2.5 Topological asymptotic expansion

In order to prove that the topological asymptotic expansion (3.4) holds, we will need the following lemmas.

Lemma 3.9 *For sufficiently small $\varepsilon \geq 0$, have the following estimation*

$$\int_{B_\varepsilon} |\nabla u_0|^2 dy = \varepsilon^2 |B| |\nabla u_0(0)|^2 + o(\varepsilon^2)$$

Proof.

To prove this lemma, it is enough to use the Taylor expansion and a change of variables. One can find the complete proof in [6]

□

Lemma 3.10 *For sufficiently small $\varepsilon \geq 0$ and $\tilde{y} = y/\varepsilon$, we have the following estimation*

$$\begin{aligned} \int_{\partial B_\varepsilon} (u_\varepsilon - u_0) \frac{\partial u_0}{\partial n} ds(y) \\ = \varepsilon^2 \nabla u_0(0) \left[\frac{1}{\alpha}(1 - \alpha) \int_{\partial B} n(\tilde{y})^T \phi(\tilde{y}) ds(\tilde{y}) \right] \nabla u_0(0)^T + o(\varepsilon^2) \end{aligned}$$

Proof.

We have

$$\begin{aligned} \int_{\partial B_\varepsilon} (u_\varepsilon - u_0) \frac{\partial u_0}{\partial n} ds(y) &= \int_{\partial B_\varepsilon} (u_\varepsilon - u_0) \nabla u_0(0) \cdot n ds(y) \\ &\quad - \int_{\partial B_\varepsilon} (u_\varepsilon - u_0) [\nabla u_0 \cdot n - \nabla u_0(0) \cdot n] ds(y) \end{aligned} \quad (3.68)$$

In [9] we can find an estimate

$$\int_{\partial B_\varepsilon} (u_\varepsilon - u_0) [\nabla u_0 \cdot n - \nabla u_0(0) \cdot n] ds(y) = o(\varepsilon^2) \quad (3.69)$$

To the first integral on the right hand side of equation (3.68) we apply a change of variables $\tilde{y} = y/\varepsilon$ and obtain

$$\begin{aligned} \int_{\partial B_\varepsilon} (u_\varepsilon - u_0) \nabla u_0(0) \cdot n ds(y) \\ = \varepsilon \int_{\partial B} [u_\varepsilon(\varepsilon \tilde{y}) - u_0(\varepsilon \tilde{y})] \nabla u_0(0) \cdot n ds(\tilde{y}) \end{aligned} \quad (3.70)$$

Applying Theorem 3.4, we get

$$\begin{aligned} \int_{\partial B_\varepsilon} (u_\varepsilon - u_0) \nabla u_0(0) \cdot n ds(y) \\ = \varepsilon^2 \nabla u_0(0) \left[\frac{1}{\alpha} (1 - \alpha) \int_{\partial B} n(\tilde{y})^T \phi(\tilde{y}) ds(\tilde{y}) \right] \nabla u_0(0)^T \end{aligned} \quad (3.71)$$

Substitution (3.69) and (3.71) into (3.68) yields to the desired result

□

Lemma 3.11 *For sufficiently small $\varepsilon \geq 0$, we have the following estimation*

$$\int_{B_\varepsilon} (u_0 - f)(u_\varepsilon - u_0) dy = o(\varepsilon^2)$$

Proof.

In the proof we follow [6]. We have

$$\int_{B_\varepsilon} (u_0 - f)(u_\varepsilon - u_0) dy \leq \|u_0 - f\|_{L^\infty(B_\varepsilon)} \int_{B_\varepsilon} |u_\varepsilon - u_0| dy$$

From the Hölder inequality, for all $p, q \in [1, \infty]$ satisfying $\frac{1}{p} + \frac{1}{q} = 1$, we obtain

$$\left| \int_{B_\varepsilon} (u_0 - f)(u_\varepsilon - u_0) dy \right| \leq C\varepsilon^{2/p} \|u_\varepsilon - u_0\|_{L^q(B_\varepsilon)}$$

We choose $p = 3/2$ and $q = 3$. It follows from the Sobolev imbedding theorem that $H^1(\Omega) \subset L^q(\Omega)$. Hence

$$\left| \int_{B_\varepsilon} (u_0 - f)(u_\varepsilon - u_0) dy \right| \leq C\varepsilon^{2/p} \|u_\varepsilon - u_0\|_{H^1(\Omega)}$$

From Lemma 3.3, we have $\|u_\varepsilon - u_0\|_{H^1(\Omega)} \leq C\varepsilon$. Thus

$$\left| \int_{B_\varepsilon} (u_0 - f)(u_\varepsilon - u_0) dy \right| \leq C\varepsilon^{7/3}$$

□

Definition 3.5 *Let the vector function ϕ be a unique solution of the problem (3.54) and n a normal vector outward to the boundary ∂B . The matrix M , defined by*

$$M = |B|I + \frac{1}{\alpha}(1 - \alpha) \int_{\partial B} n(\tilde{y})^T \phi(\tilde{y}) ds(\tilde{y}) \quad (3.72)$$

is called the polarization tensor.

The properties of the polarization tensor are investigated by Ammari et. al in [5] or in [21]. In these publications it is proven that M is a symmetric matrix. Moreover, M is positive definite for $1 > \alpha$ and it is negative definite for $1 < \alpha$. In [15] and [40] explicit form of the polarization tensor has been derived for cases where B is a ball or an ellipsoid.

Proposition 3.1 *For sufficiently small $\varepsilon \geq 0$, the following asymptotic expansion holds*

$$a_\varepsilon(u_\varepsilon, u_0) - a_0(u_\varepsilon, u_0) = \varepsilon^2(\alpha - 1) \nabla u_0(0) M \nabla u_0(0)^T + o(\varepsilon^2)$$

Proof.

From the definition of bilinear form (3.11) and the Green formula, we obtain

$$\begin{aligned}
a_\varepsilon(u_\varepsilon, u_0) - a_0(u_\varepsilon, u_0) &= (\alpha - 1) \int_{B_\varepsilon} \nabla u_\varepsilon \cdot \nabla u_0 \, dy \\
&= (\alpha - 1) \int_{B_\varepsilon} \nabla(u_\varepsilon - u_0) \cdot \nabla u_0 \, dy + (\alpha - 1) \int_{B_\varepsilon} |\nabla u_0|^2 \, dy \\
&= (\alpha - 1) \int_{\partial B_\varepsilon} (u_\varepsilon - u_0) \frac{\partial u_0}{\partial n} \, ds(y) + (\alpha - 1) \int_{B_\varepsilon} |\nabla u_0|^2 \, dy \\
&\quad - (\alpha - 1) \int_{B_\varepsilon} (u_\varepsilon - u_0)(u_0 - f) \, dy
\end{aligned}$$

To the above equation, we apply Lemma 3.9, Lemma 3.10 and Lemma 3.11. We have

$$\begin{aligned}
a_\varepsilon(u_\varepsilon, u_0) - a_0(u_\varepsilon, u_0) &= \varepsilon^2(\alpha - 1) |B| |\nabla u_0(0)|^2 \\
&\quad + \varepsilon^2(\alpha - 1) \nabla u_0(0) \left[\frac{1}{\alpha} (1 - \alpha) \int_{\partial B} n(\tilde{y})^T \phi(\tilde{y}) \, ds(\tilde{y}) \right] \nabla u_0(0)^T + o(\varepsilon^2)
\end{aligned} \tag{3.73}$$

Using Definition 3.5 we can write

$$a_\varepsilon(u_\varepsilon, u_0) - a_0(u_\varepsilon, u_0) = \varepsilon^2(\alpha - 1) \nabla u_0(0) M \nabla u_0(0)^T + o(\varepsilon^2)$$

□

Proposition 3.2 *For sufficiently small $\varepsilon \geq 0$, the following asymptotic expansion holds*

$$J_\varepsilon(u_\varepsilon) - J_0(u_0) = \varepsilon^2 \frac{(\alpha - 1)}{2} \nabla u_0(0) M \nabla u_0(0)^T + o(\varepsilon^2)$$

Proof.

Let us write the functional J_ε at u_ε with the help of bilinear a_ε and linear l form as follows

$$J_\varepsilon(u_\varepsilon) = \frac{1}{2} a_\varepsilon(u_\varepsilon, u_\varepsilon) - l(u_\varepsilon) + \frac{1}{2} \int_{\Omega} f^2 \, dy \tag{3.74}$$

Similarly, we can write the functional J_0 at u_0

$$J_0(u_0) = \frac{1}{2}a_0(u_0, u_0) - l(u_0) + \frac{1}{2} \int_{\Omega} f^2 dy \quad (3.75)$$

Subtracting (3.75) from (3.74), we get

$$J_{\varepsilon}(u_{\varepsilon}) - J_0(u_0) = \frac{1}{2}a_{\varepsilon}(u_{\varepsilon}, u_{\varepsilon}) - l(u_{\varepsilon}) - \frac{1}{2}a_0(u_0, u_0) + l(u_0)$$

Since $u_{\varepsilon} \in H^1(\Omega)$ is a unique solution of problem (3.10), we have $a_{\varepsilon}(u_{\varepsilon}, v) = l(v)$ for all $v \in H^1(\Omega)$, thus for $v = u_{\varepsilon}$ we obtain

$$a_{\varepsilon}(u_{\varepsilon}, u_{\varepsilon}) = l(u_{\varepsilon}) \quad (3.76)$$

Using the same argument as above we show that

$$a_0(u_0, u_0) = l(u_0) \quad (3.77)$$

Using (3.76) and (3.77), we can write

$$J_{\varepsilon}(u_{\varepsilon}) - J_0(u_0) = \frac{1}{2}l(u_{\varepsilon}) - l(u_{\varepsilon}) - \frac{1}{2}l(u_0) + l(u_0) = -\frac{1}{2}l(u_{\varepsilon} - u_0)$$

Since we want to avoid to calculate $u_{\varepsilon}(y) - u_0(y)$ for all $y \in \Omega$, we can write

$$J_{\varepsilon}(u_{\varepsilon}) - J_0(u_0) = -\frac{1}{2}l(u_{\varepsilon} - u_0) + \frac{1}{2}a_0(u_0, u_{\varepsilon} - u_0) - \frac{1}{2}a_0(u_0, u_{\varepsilon} - u_0)$$

where u_0 is once again the unique solution of the problem

$$\begin{cases} \text{Find } u \in H^1(\Omega) \\ a_0(u, v) = l(v) \quad \forall v \in H^1(\Omega) \end{cases} \quad (3.78)$$

Thus particularly for $v = u_{\varepsilon} - u_0$, we have

$$a_0(u_0, u_{\varepsilon} - u_0) = l(u_{\varepsilon} - u_0)$$

Using this fact, we obtain

$$\begin{aligned}
J_\varepsilon(u_\varepsilon) - J_0(u_0) &= -\frac{1}{2}a_0(u_0, u_\varepsilon - u_0) \\
&= -\frac{1}{2}a_0(u_0, u_\varepsilon) + \frac{1}{2}a_0(u_0, u_0) \\
&= -\frac{1}{2}a_0(u_\varepsilon, u_0) + \frac{1}{2}l(u_0) \\
&= -\frac{1}{2}a_0(u_\varepsilon, u_0) + \frac{1}{2}a_\varepsilon(u_\varepsilon, u_0)
\end{aligned} \tag{3.79}$$

From Proposition 3.1 we get

$$J_\varepsilon(u_\varepsilon) - J_0(u_0) = \varepsilon^2 \frac{(\alpha - 1)}{2} \nabla u_0(0) M \nabla u_0(0)^T + o(\varepsilon^2)$$

□

Corollary 3.3 *The topological gradient g is a function independent of ε , which formula is given by*

$$g(x) = \frac{(\alpha - 1)}{2} \nabla u_0(x) M \nabla u_0(x)^T \tag{3.80}$$

The symmetric and positive definite matrix M is defined as in (3.72)

Proof.

This corollary is a consequence of Proposition 3.1 and Proposition 3.2

□

3.3 Explicit formula of the topological gradient

In order to obtain an explicit formula of the topological gradient (3.80), we have to solve problem (3.54) and next to calculate the polarization tensor

$$M = |B|I + \frac{1}{\alpha}(1 - \alpha) \int_{\partial B} n(\tilde{y})^T \phi(\tilde{y}) ds \tag{3.81}$$

Using the standard approach of introducing elliptic coordinates and applying the Fourier method, Brühl et al. in [15] derived an explicit formula of the function ϕ for the case when B is a fixed ellipse, whose semi-major and semi-minor axis have length a and b , respectively, and lie on the axis of the coordinate system. This formula is as follows

$$M = |B| \begin{bmatrix} \frac{\alpha(a+b)}{(a+\alpha b)} & 0 \\ 0 & \frac{\alpha(a+b)}{(b+\alpha a)} \end{bmatrix} \quad (3.82)$$

It is not difficult to check, by setting $a = b$, that for the case when B is a ball, the matrix M has the following form

$$M = |B| \begin{bmatrix} \frac{2\alpha}{(1+\alpha)} & 0 \\ 0 & \frac{2\alpha}{(1+\alpha)} \end{bmatrix} \quad (3.83)$$

Using formulas (3.83) and (3.82) we are able to derive the formula of the topological gradient (3.80) depending on shape of the domain B

Case of B is a disk. Substituting (3.83) into (3.80) yields

$$g(x) = (\alpha - 1) \frac{\alpha}{1 + \alpha} |B| |\nabla u_0|^2 \quad (3.84)$$

Case of B is an ellipse. Using formula (3.82) we want to derive formula of topological gradient (3.80) for the case of an arbitrary ellipse. For this, we fix that $a < b$ and denote

$$m_1 = \frac{\alpha(a+b)}{(a+\alpha b)} \quad \text{and} \quad m_2 = \frac{\alpha(a+b)}{(b+\alpha a)}$$

Since

$$m_1 - m_2 = \frac{\alpha(\alpha - 1)(a^2 - b^2)}{(a + \alpha b)(b + \alpha a)} > 0 \quad \text{thus} \quad m_1 > m_2$$

Next, we assume that B_θ is an ellipse obtained by rotation of B by angle $\theta \in [0, \pi]$, then we can find an orthogonal transformation

$$R_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad (3.85)$$

such that $B_\theta = R_\theta B$. The polarization tensor corresponding to the rotated ellipse B_θ is given by

$$M_\theta = R_\theta M R_\theta^T$$

Therefore, the topological gradient for an arbitrary ellipse B_θ has the form

$$g(x, \theta) = \frac{(\alpha - 1)}{2} \nabla u_0 M_\theta \nabla u_0^T \quad (3.86)$$

We have the following proposition

Proposition 3.3 *The topological gradient $g(x, \theta)$ is minimal at the fixed point x if and only if $[\cos \theta, \sin \theta]$ is the eigenvector associated with the greatest eigenvalue of the matrix $\nabla u_0^T \nabla u_0$*

Proof.

Let us fix point $x \in \Omega$ and calculate the derivative of the topological gradient g with respect to θ . We have

$$\begin{aligned} \frac{\partial g}{\partial \theta}(x, \theta) &= (\alpha - 1)(m_1 - m_2)|B| \\ &\cdot \left(\frac{\partial u_0}{\partial x_2} \cos \theta - \frac{\partial u_0}{\partial x_1} \sin \theta \right) \left(\frac{\partial u_0}{\partial x_1} \cos \theta + \frac{\partial u_0}{\partial x_2} \sin \theta \right) \end{aligned}$$

The derivative $\frac{\partial g}{\partial \theta}(x, \theta)$ is equal to 0 if and only if

$$\theta = \arctan \left(\frac{\partial u_0}{\partial x_2} / \frac{\partial u_0}{\partial x_1} \right) =: \theta_1 \quad \text{or} \quad \theta = -\arctan \left(\frac{\partial u_0}{\partial x_1} / \frac{\partial u_0}{\partial x_2} \right) =: \theta_2$$

Let us now consider the two cases. First, let $\theta = \theta_1$. We have

$$\cos \theta_1 = \frac{1}{|\nabla u_0|} \frac{\partial u_0}{\partial x_1} \quad \text{and} \quad \sin \theta_1 = \frac{1}{|\nabla u_0|} \frac{\partial u_0}{\partial x_2}$$

Therefore

$$\frac{\partial^2 g}{\partial \theta^2}(x, \theta_1) = (\alpha - 1)(m_2 - m_1)|B||\nabla u_0|^3 > 0 \quad \text{for } m_1 > m_2$$

Next, let $\theta = \theta_2$. We have

$$\cos \theta_2 = \frac{1}{|\nabla u_0|} \frac{\partial u_0}{\partial x_2} \quad \text{and} \quad \sin \theta_2 = -\frac{1}{|\nabla u_0|} \frac{\partial u_0}{\partial x_1}$$

Therefore

$$\frac{\partial^2 g}{\partial \theta^2}(x, \theta_2) = (\alpha - 1)(m_1 - m_2)|B||\nabla u_0|^3 < 0 \quad \text{for } m_1 > m_2$$

We conclude that $g(x, \theta)$ is minimal for $\theta = \theta_1$ and

$$g(x, \theta_1) = (\alpha - 1)m_1|B||\nabla u_0|^2$$

To complete the proof, it is enough to observe that

$$\nabla u_0^T \nabla u_0 = R_{\theta_1} \begin{bmatrix} |\nabla u_0|^2 & 0 \\ 0 & 0 \end{bmatrix} R_{\theta_1}^T$$

From the above decomposition, we conclude that $\nabla u_0/|\nabla u_0|$ is the eigenvector associated with the greatest eigenvalue $|\nabla u_0|^2$ of the matrix $\nabla u_0^T \nabla u_0$

□

3.4 Remarks

As we have already mentioned, there are several approaches to derive the formula for the topological gradient and there is no uniquely justified method, which would be correct for any given cost function. In this section, we would like to give some remarks and discuss problems that arise, when we apply some of the existing methods to our particular problem, posed in Section 3.2 of the current chapter. At the end of this section, we introduce some criterion, which can be applied to edge detection. We decided to include it to the thesis, however there is no rigorous mathematical proof of its correctness. Some advantages of this criterion over magnitude of the smoothened image gradient $|\nabla u_\sigma|^2$ are confirmed based on numerical experiments, which we present in Chapter 5.

3.4.1 The generalized adjoint method

In this subsection, we would like to present a rough idea of the generalized adjoint method, introduced by C ea in [19] and further adapted to the topological sensitivity analysis by Garreau et al. in [35]. One can find a diverse modification of this method in Amstutz [6, 7].

Let $v_0 \in H^1(\Omega)$ be the unique solution of the adjoint problem

$$\begin{cases} \text{Find } v_0 \in H^1(\Omega) \\ a_0(w, v_0) = -L(w) \quad \forall w \in H^1(\Omega) \end{cases} \quad (3.87)$$

where $L : H^1(\Omega) \rightarrow \mathbb{R}$ is a continuous linear form.

Hypothesis 3.1 *For sufficiently small $\varepsilon \geq 0$, there exists a real number δJ and continuous linear form $L : H^1(\Omega) \rightarrow \mathbb{R}$, such that*

$$J_\varepsilon(u_\varepsilon) - J_0(u_0) = L(u_\varepsilon - u_0) + \varepsilon^2 \delta J + o(\varepsilon^2) \quad (3.88)$$

Hypothesis 3.2 *For sufficiently small $\varepsilon \geq 0$, there exists a real number δa , such that*

$$a_\varepsilon(u_\varepsilon, v_0) - a_0(u_\varepsilon, v_0) = \varepsilon^2 \delta a + o(\varepsilon^2)$$

We have the following proposition

Proposition 3.4 *Let u_ε be the unique solution of problem (3.10), the cost function J_ε , as defined in (3.91) and v_0 satisfies (3.87). Then, for sufficiently small $\varepsilon \geq 0$, we have the following asymptotic expansion*

$$J_\varepsilon(u_\varepsilon) - J_0(u_0) = \varepsilon^2(\delta a + \delta J) + o(\varepsilon^2)$$

Proof.

The proof is based on idea presented in [35]. From Hypothesis 3.1, we have

$$J_\varepsilon(u_\varepsilon) - J_0(u_0) = L(u_\varepsilon - u_0) + \varepsilon^2 \delta J + o(\varepsilon^2)$$

and we may write this expansion as follows

$$\begin{aligned} J_\varepsilon(u_\varepsilon) - J_0(u_0) &= L(u_\varepsilon - u_0) + a_0(u_\varepsilon - u_0, v_0) - a_0(u_\varepsilon - u_0, v_0) \\ &\quad + \varepsilon^2 \delta J + o(\varepsilon^2) \end{aligned}$$

Next, by using properties of a bilinear form, v_0 being the solution to problem (3.87) for all $w \in H^1(\Omega)$ and u_0 satisfying equation $a_0(u_0, v) = l(v)$ for all $v \in H^1(\Omega)$, we obtain

$$J_\varepsilon(u_\varepsilon) - J_0(u_0) = l(v_0) - a_0(u_\varepsilon, v_0) + \varepsilon^2 \delta J + o(\varepsilon^2)$$

Since we have assumed that the function u_ε satisfies $a_\varepsilon(u_\varepsilon, v) = l(v)$ for all $v \in H^1(\Omega)$, thus in particular for $v = v_0$. Thus, we have

$$J_\varepsilon(u_\varepsilon) - J_0(u_0) = a_\varepsilon(u_\varepsilon, v_0) - a_0(u_\varepsilon, v_0) + \varepsilon^2 \delta J + o(\varepsilon^2)$$

Finally, applying Hypothesis 3.2, we obtain

$$J_\varepsilon(u_\varepsilon) - J_0(u_0) = \varepsilon^2(\delta a + \delta J) + o(\varepsilon^2) \quad (3.89)$$

□

Therefore, if Hypothesis 3.1 and Hypothesis 3.2 would be true, we would have to find δa and δJ , in order to know the formula of the topological gradient g .

However, in Proposition 3.2 we have shown that

$$J_\varepsilon(u_\varepsilon) - J_0(u_0) = \varepsilon^2 \frac{(\alpha - 1)}{2} \nabla u_0(0) M \nabla u_0(0)^T + o(\varepsilon^2) \quad (3.90)$$

Comparing the right hand sides of equations (3.88) and (3.90) we conclude that

$$\delta J = \frac{(\alpha - 1)}{2} \nabla u_0(0) M \nabla u_0(0)^T$$

and

$$L(u_\varepsilon - u_0) = 0$$

It means, that the solution v_0 of the adjoint problem (3.87) is identically equal to 0. This implying that Hypothesis 3.2 is true only for the trivial case $\varepsilon = 0$.

3.4.2 The domain truncation method

Here, we briefly explain the idea of domain truncation method introduced by Garreau et al. in [35], further investigated by Feijóo et al. in [33]. We mention that in this subsection, we will use the same notations J_ε , a_ε , u_ε , l used in the previous parts of work, but the different problem is considered.

Let us define the cost function as follows

$$J_\varepsilon(u) := \frac{1}{2} \int_{\Omega \setminus \bar{B}_\varepsilon} (u - f)^2 dx + \frac{1}{2} \int_{\Omega \setminus \bar{B}_\varepsilon} |\nabla u|^2 dx \quad (3.91)$$

The minimum of J_ε satisfies function u_ε , solution of the problem

$$\begin{cases} u_\varepsilon - \Delta u_\varepsilon = f & x \in \Omega \setminus \bar{B}_\varepsilon \\ \frac{\partial u_\varepsilon}{\partial n} = 0 & x \in \partial B_\varepsilon \\ \frac{\partial u_\varepsilon}{\partial n} = 0 & x \in \partial \Omega \end{cases} \quad (3.92)$$

In order to apply the generalized adjoint method, we need a fixed functional space. Such a functional space, independent of ε , can be constructed by using the domain truncation technique.

Let $R > 0$ be large enough such that the open ball $B_R \subset \Omega$ includes inhomogeneity B_ε . Let us consider problems defined on the truncated domains $B_R \setminus \bar{B}_\varepsilon$ and $\Omega \setminus \bar{B}_R$. We have

$$\begin{cases} z_\varepsilon - \Delta z_\varepsilon = f & x \in B_R \setminus \bar{B}_\varepsilon \\ z_\varepsilon = \psi & x \in \partial B_R \\ \frac{\partial z_\varepsilon}{\partial n} = 0 & x \in \partial B_\varepsilon \end{cases} \quad (3.93)$$

and

$$\begin{cases} z_R - \Delta z_R = f & x \in \Omega \setminus \bar{B}_R \\ \frac{\partial z_R}{\partial n} = \Lambda_\varepsilon \psi & x \in \partial B_R \\ \frac{\partial z_R}{\partial n} = 0 & x \in \partial \Omega \end{cases} \quad (3.94)$$

where Λ_ε denotes the Steklov-Poincaré operator

$$\Lambda_\varepsilon : H^{1/2}(\partial B_R) \rightarrow H^{-1/2}(\partial B_R)$$

which maps Dirichlet data to Neumann data on the boundary ∂B_R and is defined by

$$\Lambda_\varepsilon \psi := \frac{\partial z_\varepsilon}{\partial n} \Big|_{\partial B_R}$$

We have the following proposition

Proposition 3.5 *Let $\varepsilon \geq 0$. If ψ is the trace of u_ε on ∂B_R then the restriction to $B_R \setminus \bar{B}_\varepsilon$ of the solution u_ε to problem (3.92) is the solution z_ε of problem (3.93) and the restriction to $\Omega \setminus \bar{B}_R$ of the solution u_ε is the solution z_R of problem (3.94).*

Proof.

The proof is standard, see [35]

□

From the above proposition we conclude that the function u_ε satisfies

$$\begin{cases} u_\varepsilon - \Delta u_\varepsilon = f & x \in \Omega \setminus \bar{B}_R \\ \frac{\partial u_\varepsilon}{\partial n} = \Lambda_\varepsilon u_\varepsilon & x \in \partial B_R \\ \frac{\partial u_\varepsilon}{\partial n} = 0 & x \in \partial \Omega \end{cases} \quad (3.95)$$

The bilinear form associated with the above problem is defined by

$$a_\varepsilon(u, v) := \int_{\Omega \setminus \bar{B}_R} u v \, dx + \int_{\Omega \setminus \bar{B}_R} \nabla u \cdot \nabla v \, dx - \int_{\partial B_R} \Lambda_\varepsilon u v \, ds \quad (3.96)$$

and the linear form by

$$l(v) := \int_{\Omega \setminus \bar{B}_R} f v \, dx \quad (3.97)$$

In order to prove that the topological asymptotic expansion

$$J_\varepsilon(u_\varepsilon) - J_0(u_0) = \varepsilon^2(\delta a + \delta J) + o(\varepsilon^2)$$

holds, we would have to show that for $\varepsilon \geq 0$, there exists a positive number $\delta a \in \mathbb{R}$, such that

$$a_\varepsilon(u_\varepsilon, v_0) - a_0(u_\varepsilon, v_0) = \varepsilon^2 \delta a + o(\varepsilon^2)$$

Since the domain integration does not depend on ε from definition (3.96) we obtain

$$a_\varepsilon(u_\varepsilon, v_0) - a_0(u_\varepsilon, v_0) = - \int_{\partial B_R} (\Lambda_\varepsilon - \Lambda_0) u_\varepsilon v_0 \, ds$$

For $\psi = u_\varepsilon|_{\partial B_R}$, we have

$$a_\varepsilon(u_\varepsilon, v_0) - a_0(u_\varepsilon, v_0) = - \int_{\partial B_R} \left(\frac{\partial u_\varepsilon}{\partial n} - \frac{\partial u_0}{\partial n} \right) v_0 \, ds$$

Thus, in order to find an explicit expression for δa , first we would have to solve problem (3.93), then calculate the normal derivative of its solution and last find the dominant term in the expression

$$- \int_{\partial B_R} \left(\frac{\partial u_\varepsilon}{\partial n} - \frac{\partial u_0}{\partial n} \right) v_0 \, ds$$

The natural way to deal with such a problem, defined on a ring domain, is to change the coordinate system to polar and then to apply the Fourier method. Such approach has been presented by Feijóo in [33] for the Poisson problem. In our case, since problem (3.92) is inhomogeneous, a better idea would be to use the method of Green function. However, since we would have to calculate expression $u_\varepsilon(y) - u_0(y)$ for $y \in \partial B_R$, where the radius R is fixed, we could not use asymptotic (3.26) in approximation of the fundamental solution $\Gamma(\cdot, y)$ by $\Phi(\cdot, y)$, as it was possible in (3.28).

3.4.3 Formula of the ‘topological gradient’ derived using the ‘generalized adjoint method’

Here, we propose the criterion which can be used to edge detection. Its formula is given by

$$\nabla u_\sigma \nabla v_\rho \tag{3.98}$$

The functions u_σ and v_ρ are defined by

$$u_\sigma := G_\sigma * u$$

and

$$v_\rho := G_\rho * (-\langle \nabla u_\sigma, \nabla^2 u_\sigma \nabla u_\sigma \rangle)$$

and G_σ and G_ρ denote Gaussian kernels with the standard deviation equal σ and ρ , respectively, and are defined as in (2.2).

In the definition of function v_ρ the expression

$$\langle \nabla u_\sigma, \nabla^2 u_\sigma \nabla u_\sigma \rangle = \left(\frac{\partial u_\sigma}{\partial x_2} \right)^2 \frac{\partial^2 u_\sigma}{\partial x_1^2} + 2 \frac{\partial u_\sigma}{\partial x_1} \frac{\partial u_\sigma}{\partial x_2} \frac{\partial^2 u_\sigma}{\partial x_1 \partial x_2} + \left(\frac{\partial u_\sigma}{\partial x_1} \right)^2 \frac{\partial^2 u_\sigma}{\partial x_2^2}$$

is nothing but the second order derivative of the function u_σ in the direction of ∇u_σ and is commonly used in image analysis as a second order operator for edge detection. For details we refer for instance to [46].

Chapter 4

Finite volume discretization

In this chapter, we propose the finite volume discretization for the Catté et al. (2.8) and the Weickert (2.9) models. Discretization is based on the integro-interpolation method introduced by Samarskii in [56]. The numerical schemes have been derived for the case of uniform (see Fig. 4.2) and nonuniform (see Fig. 4.3) grids for the computational domain $\Omega \subset \mathbb{R}^2$. An essential part of our discretization approach is the local adaptive coarsening strategy, which allows significant reduction of the number of grid cells.

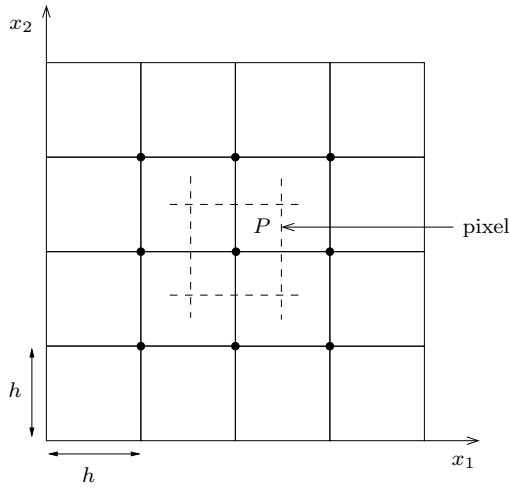


Fig. 4.1: Uniform grid on the space domain for the finite difference method

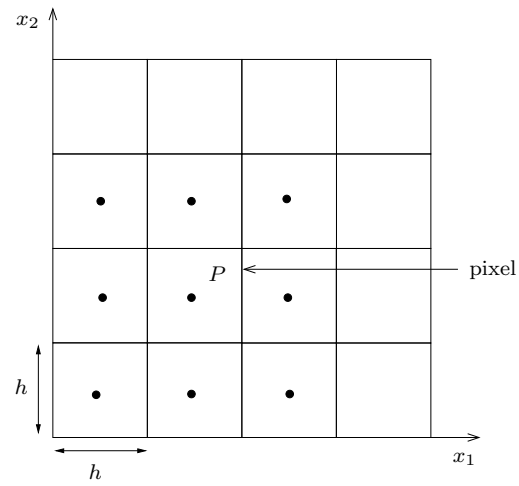


Fig. 4.2: Uniform grid on the space domain for the finite volume method

4.1 Grid coarsening technique

In this section, we describe how to construct an adaptive grid, used in discrete scale steps to solve problems (2.8) and (2.9) numerically.

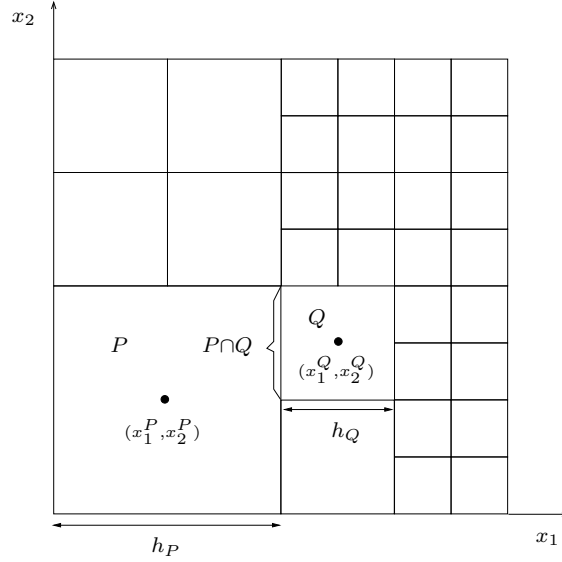


Fig. 4.3: Notations used for nonuniform grid

Let Ω_m , with $m \in \mathbb{N}$, be a uniform grid composed of square cells $P \in \Omega_m$, each with a side length $h_P = 2^{m-1}h$ and oriented along axes x_1 and x_2 . We assume that the grid points are located at the centres of cells, as presented in Fig. 4.2. We start the coarsening procedure with the fine grid Ω_1 , with grid spacing h , which is equivalent to the pixel size¹. In the whole procedure, the most important is to define the correct criterion, defining where and how to merge cells. Since the goal of the nonlinear diffusion filtering is to preserve boundaries of objects, we would like to have a fine grid at these positions. Therefore, the same criterion as in the case of edges detection can be used. In the grid coarsening procedure one can distinguish the following two steps:

Step 1. We consider union of grids Ω_1 and Ω_2 . We merge four cells, each with a side length h into one cell $Q \in \Omega_2$ with a side length $2h$ if and only if for all cells $P \in \Omega_1$, such that $P \in Q$, the criterion to edge detection at the point (x_1^P, x_2^P) is less than the fixed value of a threshold μ . The value of the function u at the central point (x_1^Q, x_2^Q) of the new cell Q is equal to the arithmetic average of values of u at the central points of cells that have been

¹In many articles from the computer vision literature is assumed that $h = 1$, which means that the pixel size is chosen as the unit of reference.

merged. As a result, we obtain a nonuniform grid composed of cells of two different sizes, h and $2h$. In general, we will denote by Ω_{m+l} a nonuniform grid composed of cells P , each with a side length $h_P \in \{2^{m-1}h, 2^mh, \dots, 2^lh\}$.

Step 2. In order to simplify the implementation process, we impose the so-called balance criterion. This implies coarsening the grid in such a way that each neighboring cell Q of the cell P will have a side length equal to h_P or two times greater or smaller than h_P . For this, we consider the union of two grids, Ω_3 and Ω_{1+2} . We merge four cells in Ω_{1+2} into one cell $Q \in \Omega_3$ if and only if for all $P \in \Omega_{1+2}$, such that $P \in Q$ and $h_P = 2h$, each neighbor of cell P has a side length equal to $2h$. We repeat this procedure until either l is equal to the required level of coarsening or 2^lh is equal to the size of image.

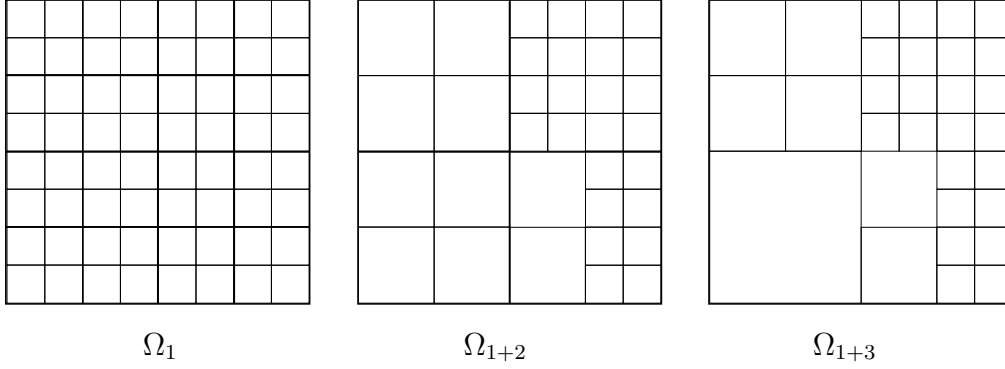


Fig. 4.4: Stages of the grid coarsening technique

The two steps described above are illustrated in Fig. 4.4. It is observed that it is not necessary to start with the fine grid Ω_1 at each iteration of the diffusion process. To save computation time, in each discrete time step, we may start the coarsening procedure with the grid obtained in the previous iteration and perform Step 1 only for the cells having a side length equal to h and then move to Step 2.

4.2 Discretization of the Catté et al. model

In this section we will derive the finite volume discretization of the Catté et al. model (2.8) using the integro-interpolation method for the case of uniform and nonuniform grid for the computational domain $\Omega \subset \mathbb{R}^2$.

Integration of the first equation in (2.8) over any finite volume $P \in \Omega_m$ or $P \in \Omega_{m+l}$, and application of the divergence theorem, yields the following integral equation

$$\int_P \frac{\partial u}{\partial t} dx = \int_{\partial P} d(|\nabla u_\sigma|^2) \frac{\partial u}{\partial n} ds \quad (4.1)$$

where n denotes the unit normal vector to the boundary ∂P , outward to P .

4.2.1 Time discretization

Let $\tau \in \mathbb{R}$ be a constant time step and $t_k := k\tau$, for $k \in \mathbb{N}$. Furthermore, let us denote $u^k(x) := u(x, t_k)$. Writing equation (4.1) at time t_k and discretizing the time partial derivative using the backward Euler method, we obtain the following equation

$$\frac{1}{\tau} \int_P (u^k - u^{k-1}) dx = \int_{\partial P} d(|\nabla u_\sigma^{k-1}|^2) \frac{\partial u^k}{\partial n} ds \quad (4.2)$$

for $k \in \mathbb{N}$ and $P \in \Omega_m$ or $P \in \Omega_{m+l}$. In order to simplify notations, we will further write d^{k-1} instead of $d(|\nabla u_\sigma^{k-1}|^2) := d(|\nabla G_\sigma * u^{k-1}(x)|)$.

Using the fact that the sides of cell P are oriented along axes x_1 and x_2 , we can write equation (4.2) in the form

$$\begin{aligned} \frac{1}{\tau} \int_P (u^k - u^{k-1}) dx &= \int_{\partial P^E} d^{k-1} \frac{\partial u^k}{\partial x_1} dx_2 - \int_{\partial P^W} d^{k-1} \frac{\partial u^k}{\partial x_1} dx_2 \\ &\quad + \int_{\partial P^N} d^{k-1} \frac{\partial u^k}{\partial x_2} dx_1 - \int_{\partial P^S} d^{k-1} \frac{\partial u^k}{\partial x_2} dx_1 \end{aligned} \quad (4.3)$$

where the upper indexes at ∂P denote respectively east, west, north and south side of the cell P .

4.2.2 Space discretization on an uniform grid

Let Ω_1 be a uniform grid of the domain $\Omega \subset \mathbb{R}^2$ composed of square cells P , each with a side length h , as presented in Fig. 4.2. We will derive a space discretization for one of the integrals on the right hand side of expression (4.3), for instance

$$\int_{\partial P^E} d^{k-1} \frac{\partial u^k}{\partial x_1} dx_2 \quad (4.4)$$

We discretize the remaining integrals in a similar way. For this, we denote the flux function by

$$W_1 = d^{k-1} \frac{\partial u^k}{\partial x_1}$$

Let us rewrite the above equation in the form

$$\frac{\partial u^k}{\partial x_1} = \frac{W_1}{d^{k-1}}$$

and then integrate this expression over the interval $[x_1^P, x_1^P + h]$

$$\int_{x_1^P}^{x_1^P+h} \frac{\partial u^k}{\partial x_1} dx_1 = \int_{x_1^P}^{x_1^P+h} \frac{W_1}{d^{k-1}} dx_1$$

Assumption that the flux function W_1 is constant on the interval $[x_1^P, x_1^P + h]$ yields

$$u(x_1^P + h, x_2) - u(x_1^P, x_2) = W_1(x_1^P + \frac{h}{2}, x_2) \int_{x_1^P}^{x_1^P+h} \frac{1}{d^{k-1}} dx_1$$

Therefore, the value of function W_1 at point $(x_1^P + \frac{h}{2}, x_2)$, where $x_2 \in [x_2^P - \frac{h}{2}, x_2^P + \frac{h}{2}]$, is given by

$$W_1(x_1^P + \frac{h}{2}, x_2) = \left[\int_{x_1^P}^{x_1^P+h} \frac{1}{d^{k-1}} dx_1 \right]^{-1} [u^k(x_1^P + h, x_2) - u^k(x_1^P, x_2)] \quad (4.5)$$

Using the midpoint rule, we can approximate the integral (4.4) as follows

$$\int_{\partial PE} d^{k-1} \frac{\partial u^k}{\partial x_1} dx_2 = \int_{\partial PE} W_1 dx_2 \approx h W_1(x_1^P + \frac{h}{2}, x_2^P) \quad (4.6)$$

Substituting (4.5) into (4.6) for $x_2 = x_2^P$ we obtain

$$\int_{\partial PE} d^{k-1} \frac{\partial u^k}{\partial x_1} dx_2 \approx h \tilde{d}^{k-1}(x_1, x_2^P) [u^k(x_1^P + h, x_2^P) - u^k(x_1^P, x_2^P)]$$

where

$$\tilde{d}^{k-1}(x_1, x_2^P) := \left[\int_{x_1^P}^{x_1^P+h} \frac{1}{d^{k-1}(x_1, x_2^P)} dx_1 \right]^{-1} \quad (4.7)$$

The function \tilde{d}^{k-1} at point $(x_1^P + \frac{h}{2}, x_2^P)$ may be approximated in one of the following two ways

$$\begin{aligned} \tilde{d}^{k-1}(x_1^P + \frac{h}{2}, x_2^P) &\approx \left[\frac{h}{d^{k-1}(x_1^P + \frac{h}{2}, x_2^P)} \right]^{-1} \\ &= \frac{d^{k-1}(x_1^P, x_2^P) + d^{k-1}(x_1^P + h, x_2^P)}{2h} \end{aligned} \quad (4.8)$$

$$\begin{aligned} \tilde{d}^{k-1}(x_1^P + \frac{h}{2}, x_2^P) &\approx \left[\frac{h}{2 d^{k-1}(x_1^P, x_2^P)} + \frac{h}{2 d^{k-1}(x_1^P + h, x_2^P)} \right]^{-1} \\ &= \frac{2 d^{k-1}(x_1^P, x_2^P) d^{k-1}(x_1^P + h, x_2^P)}{h d^{k-1}(x_1^P + h, x_2^P) + h d^{k-1}(x_1^P, x_2^P)} \end{aligned} \quad (4.9)$$

Approximation (4.8) is simply the arithmetic average of diffusivity values in the two neighboring cells and approximation (4.9) is the well known harmonic averaging. In the case of oscillatory diffusivity values, there is a suggestion in the book of Samarskii [56] to use approximation (4.9). In this book one can also find the numerical analysis of derived scheme.

4.2.3 Space discretization on a nonuniform grid

Let Ω_{m+l} be a nonuniform grid of the domain $\Omega \subset \mathbb{R}^2$ composed of square cells P , each with a side length h_P , as presented in Fig. 4.3. For every cell $P \in \Omega_{m+l}$, we consider the set of neighbors $\mathcal{N}(P)$ consisting of all cells $Q \in \Omega_{m+l}$ for which the common interface of P and Q , denoted by $P \cap Q$, has nonzero length $|P \cap Q| = \min(h_P, h_Q)$.

Let us consider discretization of the first integral on the right hand side of expression (4.3). The conservation law says that the flux through the interface ∂P^E is equal to the sum of the fluxes through interfaces ∂Q^W and ∂R^W (see Fig. 4.5). Therefore, we have

$$\int_{\partial P^E} d^{k-1} \frac{\partial u^k}{\partial x_1} dx_2 = \int_{\partial Q^W} d^{k-1} \frac{\partial u^k}{\partial x_1} dx_2 + \int_{\partial R^W} d^{k-1} \frac{\partial u^k}{\partial x_1} dx_2 \quad (4.10)$$

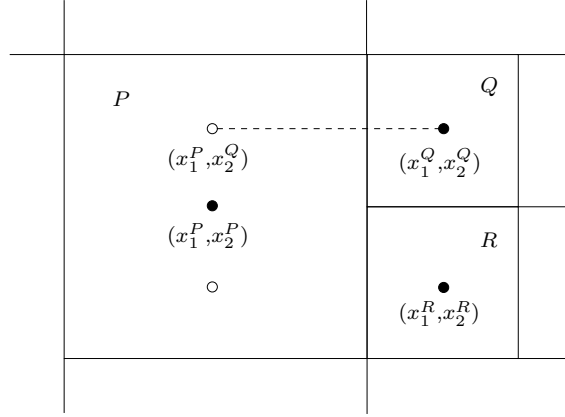


Fig. 4.5: Part of the adaptive grid

We will demonstrate how to derive a space discretization for the integral over ∂Q^W in the above expression. For that, we denote the flux function by

$$W_1 = d^{k-1} \frac{\partial u^k}{\partial x_1}$$

Let us rewrite the above equation in the form

$$\frac{\partial u^k}{\partial x_1} = \frac{W_1}{d^{k-1}}$$

and then, integrate this expression over the interval $[x_1^P, x_1^Q]$

$$\int_{x_1^P}^{x_1^Q} \frac{\partial u^k}{\partial x_1} dx_1 = \int_{x_1^P}^{x_1^Q} \frac{W_1}{d^{k-1}} dx_1$$

Assumption that the flux function W_1 is constant on the interval $[x_1^P, x_1^Q]$ yields

$$u^k(x_1^Q, x_2) - u^k(x_1^P, x_2) = W_1(x_1^P + \frac{h_P}{2}, x_2) \int_{x_1^P}^{x_1^Q} \frac{1}{d^{k-1}} dx_1$$

Therefore, the value of function W_1 at point $(x_1^P + \frac{h_P}{2}, x_2)$, where $x_2 \in [x_2^P, x_2^P + \frac{h_P}{2}]$, is given by

$$W_1(x_1^P + \frac{h_P}{2}, x_2) = \left[\int_{x_1^P}^{x_1^Q} \frac{1}{d^{k-1}} dx_1 \right]^{-1} [u^k(x_1^Q, x_2) - u^k(x_1^P, x_2)] \quad (4.11)$$

Using the midpoint rule, we approximate the integral over ∂Q^W in expression (4.10) as follows

$$\int_{\partial Q^W} d^{k-1} \frac{\partial u}{\partial x_1} dx_2 = \int_{\partial Q^W} W_1 dx_2 \approx h_Q W_1(x_1^P + \frac{h_P}{2}, x_2^Q) \quad (4.12)$$

Substituting (4.31) into (4.33) for $x_2 = x_2^P$, we obtain

$$\int_{\partial Q^W} d^{k-1} \frac{\partial u^k}{\partial x_1} dx_2 \approx \tilde{d}^{k-1}(x_1, x_2^Q) h_Q [u^k(x_1^Q, x_2^Q) - u^k(x_1^P, x_2^Q)]$$

where

$$\tilde{d}^{k-1}(x_1, x_2^Q) := \left[\int_{x_1^P}^{x_1^Q} \frac{1}{d^{k-1}(x_1, x_2^Q)} dx_1 \right]^{-1} \quad (4.13)$$

The function \tilde{d}^{k-1} at the point $(x_1^P + \frac{h}{2}, x_2^Q)$ may be approximated in one of the following two ways

$$\begin{aligned} \tilde{d}^{k-1}(x_1^P + \frac{h}{2}, x_2^Q) &\approx \left[\frac{h_P + h_Q}{2 d^{k-1}(x_1^P + \frac{h}{2}, x_2^Q)} \right]^{-1} \\ &= \frac{d^{k-1}(x_1^P, x_2^Q) + d^{k-1}(x_1^Q, x_2^Q)}{h_P + h_Q} \end{aligned} \quad (4.14)$$

$$\begin{aligned} \tilde{d}^{k-1}(x_1^P + \frac{h}{2}, x_2^Q) &\approx \left[\frac{h_P}{2 d^{k-1}(x_1^P, x_2^Q)} + \frac{h_Q}{2 d^{k-1}(x_1^Q, x_2^Q)} \right]^{-1} \\ &= \frac{2 d^{k-1}(x_1^P, x_2^Q) d^{k-1}(x_1^Q, x_2^Q)}{h_P d^{k-1}(x_1^Q, x_2^Q) + h_Q d^{k-1}(x_1^P, x_2^Q)} \end{aligned} \quad (4.15)$$

Since point (x_1^P, x_2^Q) , marked by \circ in Fig. 4.5, is not a grid point, values of functions u^k and d^{k-1} at this point have to be interpolated. We use the

simplest constant interpolation, it means, we assume that values of u^k and d^{k-1} at the point (x_1^P, x_2^Q) are equal to values of these functions at the point (x_1^P, x_2^P) .

In [30], Ewing et al. proposed a finite difference method on rectangular cell-centered grids with local refinement in order to derive discretization of the second-order elliptic equation. They have proven that the discretization for the case of a constant interpolation of irregular points provides simple symmetric scheme with error estimates in the discrete H^1 -norm of order $h^{1/2}$. In the case of a nonsymmetric and a more accurate symmetric scheme, the convergence rate is $O(h^{3/2})$.

4.3 Discretization of the Weickert model

In this section we will derive the finite volume discretization of the Weickert model (2.9) using the integro-interpolation method for the case of uniform and nonuniform grid of the computational domain $\Omega \subset \mathbb{R}^2$.

Integration of the first equation in (2.9) over the finite volume $P \subset \Omega_m$ or $P \in \Omega_{m+l}$ and application of the divergence theorem yields the following balance equation

$$\int_P \frac{\partial u}{\partial t} dx = \int_{\partial P} \langle D(S_\rho(\nabla u_\sigma)) \nabla u, n \rangle ds \quad (4.16)$$

where n denotes the unit normal vector to the boundary ∂P , outward to P and the diffusion tensor $D(S_\rho(\nabla u_\sigma))$ is defined as in Section 2.3 of Chapter 2.

4.3.1 Time discretization

Let $\tau \in \mathbb{R}$ be a constant time step and $t_k := k\tau$, for $k \in \mathbb{N}$. Furthermore, let us denote $u^k(x) := u(x, t_k)$. Writing equation (4.16) at time t_k and discretizing the time partial derivative by the Euler explicit scheme, we obtain the following equation

$$\frac{1}{\tau} \int_P (u^k - u^{k-1}) dx = \int_{\partial P} \langle D(S_\rho(\nabla u_\sigma^{k-1})) \nabla u^k, n \rangle ds \quad (4.17)$$

for all $P \in \Omega_m$ or $P \in \Omega_{m+l}$. Further, we denote entries of the diffusion tensor $D(S_\rho(\nabla u_\sigma^{k-1}))$ by d_{ij}^{k-1} , where $i, j = 1, 2$.

Using the fact that the sides of cell P are oriented along axes x_1 and x_2 , we can write equation (4.17) in the form

$$\begin{aligned} & \frac{1}{\tau} \int_P (u^k - u^{k-1}) dx = \\ & \int_{\partial P^E} \left(d_{11}^{k-1} \frac{\partial u^k}{\partial x_1} + d_{12}^{k-1} \frac{\partial u^k}{\partial x_2} \right) dx_2 - \int_{\partial P^W} \left(d_{11}^{k-1} \frac{\partial u^k}{\partial x_1} + d_{12}^{k-1} \frac{\partial u^k}{\partial x_2} \right) dx_2 \\ & + \int_{\partial P^N} \left(d_{21}^{k-1} \frac{\partial u^k}{\partial x_1} + d_{22}^{k-1} \frac{\partial u^k}{\partial x_2} \right) dx_1 - \int_{\partial P^S} \left(d_{21}^{k-1} \frac{\partial u^k}{\partial x_1} + d_{22}^{k-1} \frac{\partial u^k}{\partial x_2} \right) dx_1 \end{aligned} \quad (4.18)$$

The upper indexes at ∂P denote the east, west, north and south side of the cell P , respectively.

4.3.2 Space discretization on an uniform grid

Let Ω_1 be a uniform grid of the domain $\Omega \subset \mathbb{R}^2$ composed of square cells $P \in \Omega_1$, each with a side length h , as it is presented in Fig. 4.2. We will derive a space discretization for one of the integrals on the right hand side of expression (4.18), for example

$$\int_{\partial P^E} \left(d_{11}^{k-1} \frac{\partial u^k}{\partial x_1} + d_{12}^{k-1} \frac{\partial u^k}{\partial x_2} \right) dx_2 \quad (4.19)$$

We discretize the integrals in a similar way. For this, the flux function is denoted by

$$W_2 = d_{11}^{k-1} \frac{\partial u^k}{\partial x_1} + d_{12}^{k-1} \frac{\partial u^k}{\partial x_2}$$

Let us rewrite the above equation in the form

$$\frac{\partial u^k}{\partial x_1} + \frac{d_{12}^{k-1}}{d_{11}^{k-1}} \frac{\partial u^k}{\partial x_2} = \frac{W_2}{d_{11}^{k-1}}$$

and then integrate this expression over the interval $[x_1^P, x_1^P + h]$

$$\int_{x_1^P}^{x_1^P + h} \left(\frac{\partial u^k}{\partial x_1} + \frac{d_{12}^{k-1}}{d_{11}^{k-1}} \frac{\partial u^k}{\partial x_2} \right) dx_1 = \int_{x_1^P}^{x_1^P + h} \frac{W_2}{d_{11}^{k-1}} dx_1 \quad (4.20)$$

Assumption that the flux function W_2 is constant on the interval $[x_1^P, x_1^P + h]$ yields

$$\int_{x_1^P}^{x_1^P+h} \frac{W_2}{d_{11}^{k-1}} dx_1 = W_2(x_1^P + \frac{h}{2}, x_2) \int_{x_1^P}^{x_1^P+h} \frac{1}{d_{11}^{k-1}} dx_1 \quad (4.21)$$

We approximate the left hand side of equation (4.20) as follows

$$\begin{aligned} \int_{x_1^P}^{x_1^P+h} \left(\frac{\partial u^k}{\partial x_1} + \frac{d_{12}^{k-1}}{d_{11}^{k-1}} \frac{\partial u^k}{\partial x_2} \right) dx_1 &\approx u^k(x_1^P + h, x_2) - u^k(x_1^P, x_2) \\ &+ d_{12}^{k-1}(x_1^P + \frac{h}{2}, x_2) \frac{\partial u^k}{\partial x_2}(x_1^P + \frac{h}{2}, x_2) \int_{x_1^P}^{x_1^P+h} \frac{1}{d_{11}^{k-1}} dx_1 \end{aligned} \quad (4.22)$$

By substituting (4.21) and (4.22) into equation (4.20), an approximate value of the function W_2 at point $(x_1^P + \frac{h}{2}, x_2)$ is given by

$$\begin{aligned} W_2(x_1^P + \frac{h}{2}, x_2) &\approx \left[\int_{x_1^P}^{x_1^P+h} \frac{1}{d_{11}^{k-1}} dx_1 \right]^{-1} [u^k(x_1^P + h, x_2) - u^k(x_1^P, x_2)] \\ &+ d_{12}^{k-1}(x_1^P + \frac{h}{2}, x_2) \frac{\partial u^k}{\partial x_2}(x_1^P + \frac{h}{2}, x_2) \end{aligned} \quad (4.23)$$

where $x_2 \in [x_2^P - \frac{h}{2}, x_2^P + \frac{h}{2}]$. We further approximate the last term in (4.23) by the arithmetic mean as follows

$$\begin{aligned} &d_{12}^{k-1}(x_1^P + \frac{h}{2}, x_2) \frac{\partial u^k}{\partial x_2}(x_1^P + \frac{h}{2}, x_2) \\ &\approx \frac{1}{2} \left[d_{12}^{k-1}(x_1^P, x_2) \frac{\partial u^k}{\partial x_2}(x_1^P, x_2) + d_{12}^{k-1}(x_1^P + h, x_2) \frac{\partial u^k}{\partial x_2}(x_1^P + h, x_2) \right] \end{aligned} \quad (4.24)$$

Application of the midpoint rule to integral (4.19) yields

$$\int_{\partial P^E} \left(d_{11}^{k-1} \frac{\partial u}{\partial x_1} + d_{12}^{k-1} \frac{\partial u^k}{\partial x_2} \right) dx_2 = \int_{\partial P^E} W_2 dx_2 \approx h W_2(x_1^P + \frac{h}{2}, x_2^P) \quad (4.25)$$

By substituting (4.23) into (4.25), we obtain for $x_2 = x_2^P$

$$\begin{aligned}
\int_{\partial P^E} \left(d_{11}^{k-1} \frac{\partial u^k}{\partial x_1} + d_{12}^{k-1} \frac{\partial u^k}{\partial x_2} \right) dx_2 &\approx \tilde{d}_{11}^{k-1}(x_1, x_2^P) h [u^k(x_1^P + h, x_2) - u^k(x_1^P, x_2)] \\
&+ \frac{h}{2} \left[d_{12}^{k-1}(x_1^P, x_2) \frac{\partial u^k}{\partial x_2}(x_1^P, x_2) + d_{12}^{k-1}(x_1^P + h, x_2) \frac{\partial u^k}{\partial x_2}(x_1^P + h, x_2) \right]
\end{aligned} \tag{4.26}$$

where

$$\tilde{d}_{11}^{k-1}(x_1, x_2^P) := \left[\int_{x_1^P}^{x_1^P + h} \frac{1}{d_{11}^{k-1}(x_1, x_2^P)} dx_1 \right]^{-1}$$

The partial derivatives on the right hand side of expression (4.26) at points (x_1^P, x_2^P) and $(x_1^P + h, x_2^P)$ may be approximated using the central difference

$$\begin{aligned}
\frac{\partial u}{\partial x_2}(x_1^P, x_2^P) &\approx \frac{1}{2h} [u^k(x_1^P, x_2^P + h) - u^k(x_1^P, x_2^P - h)] \\
\frac{\partial u}{\partial x_2}(x_1^P + h, x_2^P) &\approx \frac{1}{2h} [u^k(x_1^P + h, x_2^P + h) - u^k(x_1^P + h, x_2^P - h)]
\end{aligned}$$

The function \tilde{d}_{11}^{k-1} at the point $(x_1^P + \frac{h}{2}, x_2^P)$ may be approximated in a similar way as the function \tilde{d}^{k-1} in (4.8) or in (4.9).

Notice that if we would derive, in the same manner as above, an approximation of the integral over ∂P^W , appearing on the right hand side of equation (4.18), and next the result subtract from (4.26), the term

$$d_{12}^{k-1}(x_1^P, x_2) \frac{\partial u^k}{\partial x_2}(x_1^P, x_2)$$

would cancel out. The derived space discretization of the operator $\nabla \cdot (D \nabla u)$ is equivalent to what Samarskii proposed in [56, page 288] and provide an approximation of $O(h^2)$.

4.3.3 Space discretization on a nonuniform grid

Let Ω_{m+l} be a nonuniform grid of domain $\Omega \subset \mathbb{R}^2$ composed of square cells P , each with a side length h_P , as presented in Fig. 4.3. For every cell $P \in \Omega_{m+l}$, we consider the set of neighbors $\mathcal{N}(P)$ consisting of all cells $Q \in \Omega_{m+l}$ for

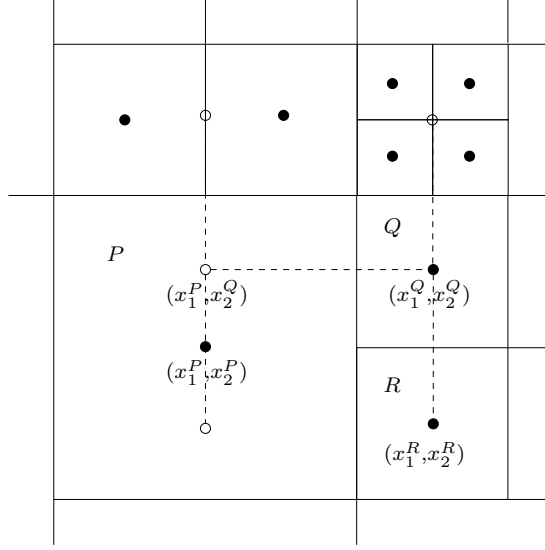


Fig. 4.6: Part of the adaptive grid

which the common interface of P and Q , denoted by $P \cap Q$, has nonzero length $|P \cap Q| = \min(h_P, h_Q)$.

Let us consider the discretization of the first integral on the right hand side of expression (4.18). The conservation law says that the flux through the interface ∂P^E is equal to the sum of the fluxes through interfaces ∂Q^W and ∂R^W (see Fig. 4.6). Therefore, we have

$$\begin{aligned} \int_{\partial P^E} \left(d_{11}^{k-1} \frac{\partial u^k}{\partial x_1} + d_{12}^{k-1} \frac{\partial u^k}{\partial x_2} \right) dx_2 &= \int_{\partial Q^W} \left(d_{11}^{k-1} \frac{\partial u^k}{\partial x_1} + d_{12}^{k-1} \frac{\partial u^k}{\partial x_2} \right) dx_2 \\ &\quad + \int_{\partial R^W} \left(d_{11}^{k-1} \frac{\partial u^k}{\partial x_1} + d_{12}^{k-1} \frac{\partial u^k}{\partial x_2} \right) dx_2 \end{aligned} \quad (4.27)$$

We will show the derivation of a space discretization for the integral over ∂Q^W . For this, we denote the flux function by

$$W_2 = d_{11}^{k-1} \frac{\partial u^k}{\partial x_1} + d_{12}^{k-1} \frac{\partial u^k}{\partial x_2}$$

Let us rewrite the above equation in the form

$$\frac{\partial u^k}{\partial x_1} + \frac{d_{12}^{k-1}}{d_{11}^{k-1}} \frac{\partial u^k}{\partial x_2} = \frac{W_2}{d_{11}^{k-1}}$$

and then, integrate this expression over the interval $[x_1^P, x_1^Q]$

$$\int_{x_1^P}^{x_1^Q} \left(\frac{\partial u^k}{\partial x_1} + \frac{d_{12}^{k-1}}{d_{11}^{k-1}} \frac{\partial u^k}{\partial x_2} \right) dx_1 = \int_{x_1^P}^{x_1^Q} \frac{W_2}{d_{11}^{k-1}} dx_1 \quad (4.28)$$

Assumption that the flux function W_2 is constant on the interval $[x_1^P, x_1^Q]$ yields

$$\int_{x_1^P}^{x_1^Q} \frac{W_2}{d_{11}^{k-1}} dx_1 = W_2(x_1^P + \frac{h_P}{2}, x_2) \int_{x_1^P}^{x_1^Q} \frac{1}{d_{11}^{k-1}} dx_1 \quad (4.29)$$

We approximate the left hand side of equation (4.28) as follows

$$\begin{aligned} \int_{x_1^P}^{x_1^Q} \left(\frac{\partial u^k}{\partial x_1} + \frac{d_{12}^{k-1}}{d_{11}^{k-1}} \frac{\partial u^k}{\partial x_2} \right) dx_1 &\approx u^k(x_1^Q, x_2) - u^k(x_1^P, x_2) \\ &+ d_{12}^{k-1}(x_1^P + \frac{h_P}{2}, x_2) \frac{\partial u^k}{\partial x_2}(x_1^P + \frac{h_P}{2}, x_2) \int_{x_1^P}^{x_1^Q} \frac{1}{d_{11}^{k-1}} dx_1 \end{aligned} \quad (4.30)$$

By substituting (4.29) and (4.30) into equation (4.28), an approximate value of the function W_2 at point $(x_1^P + \frac{h_P}{2}, x_2)$ is given by

$$\begin{aligned} W_2(x_1^P + \frac{h_P}{2}, x_2) &\approx \left[\int_{x_1^P}^{x_1^Q} \frac{1}{d_{11}^{k-1}} dx_1 \right]^{-1} [u^k(x_1^Q, x_2) - u^k(x_1^P, x_2)] \\ &+ d_{12}^{k-1}(x_1^P + \frac{h_P}{2}, x_2) \frac{\partial u^k}{\partial x_2}(x_1^P + \frac{h_P}{2}, x_2) \end{aligned} \quad (4.31)$$

where $x_2 \in [x_2^P, x_2^P + \frac{h_P}{2}]$. The last term in (4.31) is approximated by the weighted mean as follows

$$\begin{aligned} d_{12}^{k-1}(x_1^P + \frac{h}{2}, x_2) \frac{\partial u^k}{\partial x_2}(x_1^P + \frac{h}{2}, x_2) &\approx \frac{h_P}{h_P + h_Q} d_{12}^{k-1}(x_1^P, x_2) \frac{\partial u^k}{\partial x_2}(x_1^P, x_2) \\ &+ \frac{h_Q}{h_P + h_Q} d_{12}^{k-1}(x_1^Q, x_2) \frac{\partial u^k}{\partial x_2}(x_1^Q, x_2) \end{aligned} \quad (4.32)$$

The integral over ∂Q^W in expression (4.27) may be approximated using the midpoint rule as follows

$$\begin{aligned} \int_{\partial Q^W} \left(d_{11}^{k-1} \frac{\partial u^k}{\partial x_1} + d_{12}^{k-1} \frac{\partial u^k}{\partial x_2} \right) dx_2 &= \int_{\partial Q^W} W_2 dx_2 \\ &\approx h_Q W_2(x_1^P + \frac{h_P}{2}, x_2^Q) \end{aligned} \quad (4.33)$$

By substituting (4.31) into (4.33), we obtain for $x_2 = x_2^Q$

$$\begin{aligned} \int_{\partial Q^W} \left(d_{11}^{k-1} \frac{\partial u^k}{\partial x_1} + d_{12}^{k-1} \frac{\partial u^k}{\partial x_2} \right) dx_2 &\approx \tilde{d}_{11}^{k-1}(x_1, x_2^Q) h_Q [u^k(x_1^Q, x_2^Q) - u^k(x_1^P, x_2^Q)] \\ &+ h_Q \left[\frac{h_P}{h_P + h_Q} d_{12}^{k-1}(x_1^P, x_2^Q) \frac{\partial u^k}{\partial x_2}(x_1^P, x_2^Q) + \frac{h_P}{h_P + h_Q} d_{12}^{k-1}(x_1^Q, x_2^Q) \frac{\partial u^k}{\partial x_2}(x_1^Q, x_2^Q) \right] \end{aligned}$$

The partial derivatives on the right hand side of the above expression may be approximated by using the midpoint rule as follows

$$\begin{aligned} \frac{\partial u^k}{\partial x_2}(x_1^P, x_2^Q) &\approx \frac{1}{h_P} [u^k(x_1^P, x_2^Q + h_Q) - u^k(x_1^P, x_2^Q - h_Q)] \\ \frac{\partial u^k}{\partial x_2}(x_1^Q, x_2^Q) &\approx \frac{1}{h_P} [u^k(x_1^Q, x_2^Q + h_Q) - u^k(x_1^Q, x_2^Q - h_Q)] \end{aligned}$$

The function \tilde{d}_{11}^{k-1} is defined by

$$\tilde{d}_{11}^{k-1}(x_1, x_2^Q) := \left[\int_{x_1^P}^{x_1^Q} \frac{1}{d_{11}^{k-1}(x_1, x_2^Q)} dx_1 \right]^{-1}$$

At the point $(x_1^P + \frac{h_P}{2}, x_2^Q)$, it could be approximated similar to the function \tilde{d}^{k-1} in (4.14) or in (4.15).

The points, marked by \circ in Fig. 4.6, are not grid points and values of the function u^k at these points have to be interpolated. We consider a constant interpolation in the cell P , which means that values of u^k at points (x_1^P, x_2^Q) and (x_1^P, x_2^R) are equal to $u^k(x_1^P, x_2^P)$. The value of u^k at the point $(x_1^P, x_2^Q + h_Q)$ is equal to the arithmetic mean of values u^k in two neighboring cells. The value of u^k at the point $(x_1^Q, x_2^Q + h_Q)$ is equal to the arithmetic mean of values u^k in four neighboring cells.

4.4 Discretization of the first and the second order derivatives

In this section, we present how to approximate the magnitude of gradient ∇u_σ

$$|\nabla u_\sigma|^2 = \left(\frac{\partial u_\sigma}{\partial x_1} \right)^2 + \left(\frac{\partial u_\sigma}{\partial x_2} \right)^2 \quad (4.34)$$

and the second order derivative of the function u_σ in the gradient ∇u_σ direction

$$\begin{aligned} & \langle \nabla u_\sigma, \nabla^2 u_\sigma \nabla u_\sigma \rangle \\ &= \left(\frac{\partial u_\sigma}{\partial x_2} \right)^2 \frac{\partial^2 u_\sigma}{\partial x_1^2} + 2 \frac{\partial u_\sigma}{\partial x_1} \frac{\partial u_\sigma}{\partial x_2} \frac{\partial^2 u_\sigma}{\partial x_1 \partial x_2} + \left(\frac{\partial u_\sigma}{\partial x_1} \right)^2 \frac{\partial^2 u_\sigma}{\partial x_2^2} \end{aligned} \quad (4.35)$$

on an uniform as well as a nonuniform grid.

4.4.1 Discretization on an uniform grid

To approximate values of expressions (4.34) and (4.35) for an uniform grid we use standard second order approximation. The first derivative of the function u_σ is approximated using the central difference scheme

$$\frac{\partial u_\sigma}{\partial x_1}(x_1^P, x_2^P) \approx \frac{1}{2h} [u_\sigma(x_1^P + h, x_2^P) - u_\sigma(x_1^P - h, x_2^P)]$$

The second derivative is approximated by

$$\begin{aligned} \frac{\partial^2 u_\sigma}{\partial x_1^2}(x_1^P, x_2^P) &\approx \frac{1}{h} \left[\frac{\partial u_\sigma}{\partial x_1}(x_1^P + \frac{h}{2}, x_2^P) - \frac{\partial u_\sigma}{\partial x_1}(x_1^P - \frac{h}{2}, x_2^P) \right] \\ &\approx \frac{1}{h} [u_\sigma(x_1^P + h, x_2^P) - 2u_\sigma(x_1^P, x_2^P) + u_\sigma(x_1^P - h, x_2^P)] \end{aligned}$$

The mixed derivative of the function u_σ is approximated by

$$\frac{\partial^2 u_\sigma}{\partial x_1 \partial x_2}(x_1^P, x_2^P) \approx \frac{1}{2h} \left[\frac{\partial u_\sigma}{\partial x_1}(x_1^P, x_2^P + h) - \frac{\partial u_\sigma}{\partial x_1}(x_1^P, x_2^P - h) \right]$$

4.4.2 Discretization on a nonuniform grid

Since it is difficult to write discretization for the general case of a nonuniform grid, we will demonstrate how to approximate values of expressions (4.34) and (4.35) for the configuration of cells presented in Fig. 4.7

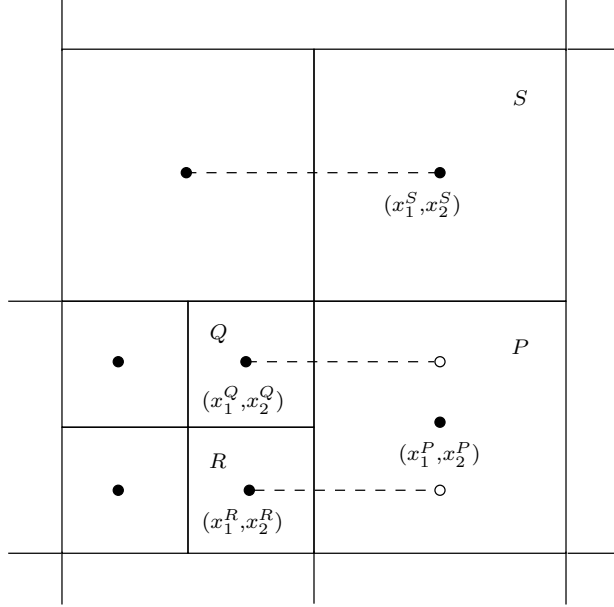


Fig. 4.7: Part of the adaptive grid

In order to compute expression (4.34) and (4.35), we have to know values of the first and the second order derivatives of the function u_σ at point (x_1^P, x_2^P) , the center of the cell P . The first derivative is approximated by the arithmetic mean of the derivatives at points $(x_1^P - \frac{h_P}{2}, x_2^P)$ and $(x_1^P + \frac{h_P}{2}, x_2^P)$ as follows

$$\frac{\partial u_\sigma}{\partial x_1}(x_1^P, x_2^P) \approx \frac{1}{2} \left[\frac{\partial u_\sigma}{\partial x_1}(x_1^P + \frac{h_P}{2}, x_2^P) + \frac{\partial u_\sigma}{\partial x_1}(x_1^P - \frac{h_P}{2}, x_2^P) \right]$$

The second derivative is approximated by the central difference of derivatives at these points

$$\frac{\partial^2 u_\sigma}{\partial x_1^2}(x_1^P, x_2^P) \approx \frac{1}{h_P} \left[\frac{\partial u_\sigma}{\partial x_1}(x_1^P + \frac{h_P}{2}, x_2^P) - \frac{\partial u_\sigma}{\partial x_1}(x_1^P - \frac{h_P}{2}, x_2^P) \right]$$

The derivative at point $(x_1^P - \frac{h_P}{2}, x_2^P)$ may be approximated as follows

$$\frac{\partial u_\sigma}{\partial x_1}(x_1^P - \frac{h_P}{2}, x_2^P) \approx \frac{1}{2} \left[\frac{u_\sigma(x_1^P, x_2^Q) - u_\sigma(x_1^Q, x_2^Q)}{\frac{1}{2}(h_P + h_Q)} + \frac{u_\sigma(x_1^P, x_2^R) - u_\sigma(x_1^R, x_2^R)}{\frac{1}{2}(h_P + h_R)} \right]$$

Since (x_1^P, x_2^Q) and (x_1^P, x_2^R) are not the grid points, the values of the function u at these points have to be interpolated. For that, we use simple constant interpolation, it means, we assume that these values are equal to value of function u at the point (x_1^P, x_2^P) .

The value of the mixed derivative at point (x_1^P, x_2^P) is approximated by the arithmetic mean of the mixed derivative values at all corners of the cell P . We demonstrate how to approximate the value of the mixed derivative in the upper left corner of the cell P

$$\frac{\partial^2 u_\sigma}{\partial x_1 \partial x_2}(x_1^P - \frac{h_P}{2}, x_2^P + \frac{h_P}{2}) \approx \frac{2}{h_P + h_S} \left[\frac{\partial u_\sigma}{\partial x_1}(x_1^S - \frac{h_S}{2}, x_2^S) - \frac{\partial u_\sigma}{\partial x_1}(x_1^P - \frac{h_P}{2}, x_2^P) \right]$$

In a similar manner, we approximate the first and the second order partial derivatives of a function u_σ with respect to the variable x_2 .

Chapter 5

Numerical experiments

In this chapter, we perform experiments which aims to compare the criterion $\nabla u_\sigma \nabla v_\rho$ that was introduced in Section 3.4 of Chapter 3, with the commonly used magnitude of the smoothed image gradient $|\nabla u_\sigma|^2$ in application to edge detection. In the second part, we present results of nonadaptive and adaptive diffusion filtering using semi-implicit finite volume schemes derived in Section 4.2 of Chapter 4.

Experiments presented in this chapter have been performed with some real and artificial images¹. In all examples, we work with gray pictures with intensity between 0 and 255. For implementation, we used Matlab ver. 5.3.

5.1 Edge detection

Let us recall definitions of function u_σ and function v_ρ . We have

$$u_\sigma := G_\sigma * u$$

and

$$v_\rho := G_\rho * (-\langle \nabla u_\sigma, \nabla^2 u_\sigma \nabla u_\sigma \rangle)$$

where G_σ and G_ρ denote Gaussian kernel with the standard deviation equal to σ and to ρ , respectively.

In order to roughly illustrate the difference between criterion $|\nabla u_\sigma|^2$ and $\nabla u_\sigma \nabla v_\rho$, we model one dimensional edge with a smooth function $u(x) = 1 + \tanh(10x)$. The plot of this function is presented in Fig. 5.1.

¹The test images presented in Fig. 5.12 and Fig. 5.16 are taken from the Institute of Computer Science, University of Innsbruck.

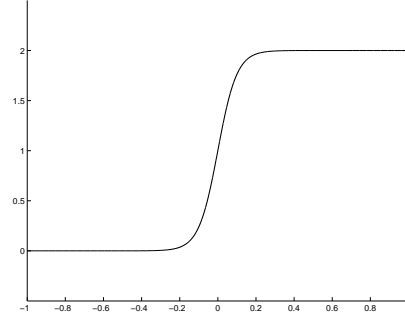
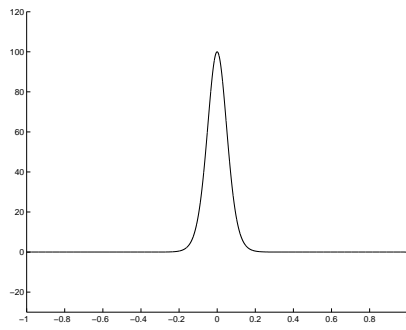
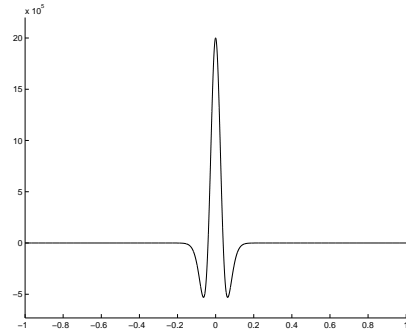


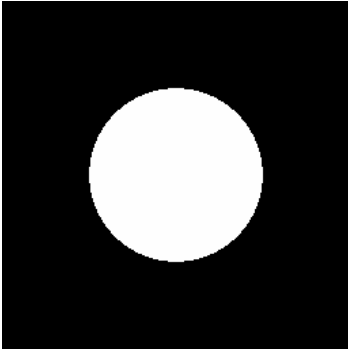
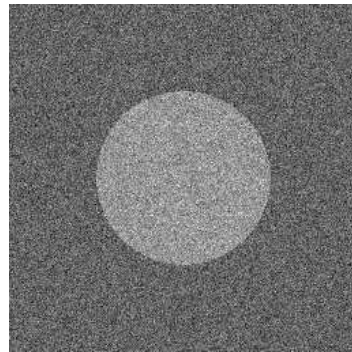
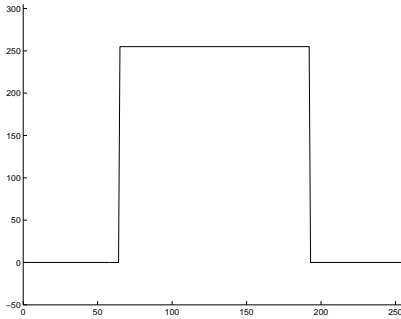
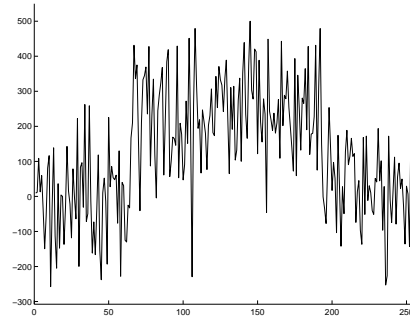
Fig. 5.1: Simple edge model

For simplicity, we neglect smoothing by a convolution with the Gaussian kernel. Therefore, we can now easily derive explicit formulas for both criteria. Its plots are presented in Fig. 5.2 and in Fig. 5.3. As we expected, in the both cases the maxima of criterion indicate the center of our model edge. However, in the case of real images, the problem of edge detection is more complicated because of many local maxima. In order to approximate a position of the edge points, we have to apply the threshold operator. This implies that we have to determine some value for μ , such that for all points x , at which the criterion is greater than μ , will lead to the set of edge points. Of course, when choosing the value for μ , we should remember that the lower a threshold, implies that more lines will be detected, and the results become increasingly susceptible to noise, and also to picking out irrelevant features from the image. Conversely, a high threshold may ignore subtle or segmented lines.

Fig. 5.2: Plot of $(u')^2$ Fig. 5.3: Plot of $u'v'$

In the next part of this chapter, we present results of the three experiments in order to compare the criteria $\nabla u_\sigma \nabla v_\rho$ and $|\nabla u_\sigma|^2$ in application to edge detection for the case of real images.

In the first experiment, we would like to compare behavior of the both criteria with respect to the presence of noise in an image. In Fig. 5.5, we see the true image f perturbed by the additive noise with the normal distribution of mean 0 and standard deviation $255/2$. In order to better illustrate the problem of edge detection, we restrict ourself to the cross-sections of considered images passing through its centers, although all calculations have been performed for the two dimensional case.

Fig. 5.4: True image f Fig. 5.5: Noisy image u Fig. 5.6: Cross-section of f Fig. 5.7: Cross-section of u

To calculate the criteria $|\nabla u_\sigma|^2$ and $\nabla u_\sigma \nabla v_\rho$ for a noisy image u , we applied a smoothing by convolution with the Gaussian kernel with parameters $\sigma = \rho = 1$. Plots in Fig. 5.8 and in Fig. 5.9 present cross-sections of the both criteria and histograms are shown in Fig. 5.10 and in Fig. 5.11. The first thing that should be noticed is that the criterion $\nabla u_\sigma \nabla v_\rho$ is less sensitive to large oscillations of image values that were introduced by noise. Due to this, it is easier to choose the right threshold value without picking up irrelevant edges.

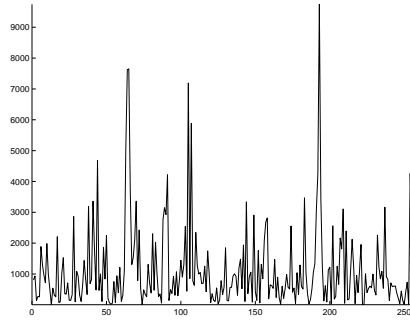


Fig. 5.8: Cross-section of $|\nabla u_\sigma|^2$ with $\sigma = 1$

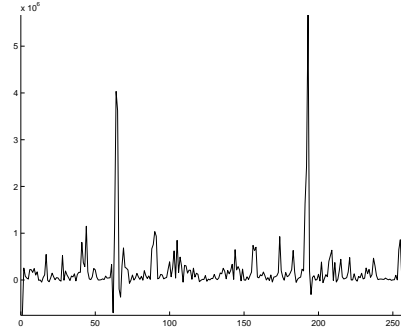


Fig. 5.9: Cross-section of $\nabla u_\sigma \nabla v_\rho$ with $\sigma = \rho = 1$

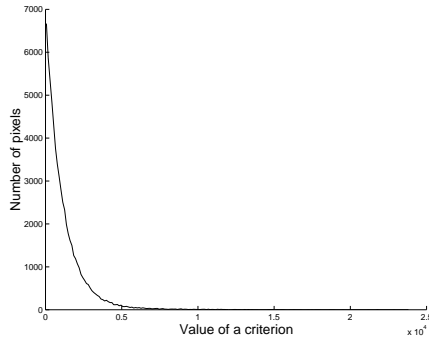


Fig. 5.10: Histogram of $|\nabla u_\sigma|^2$

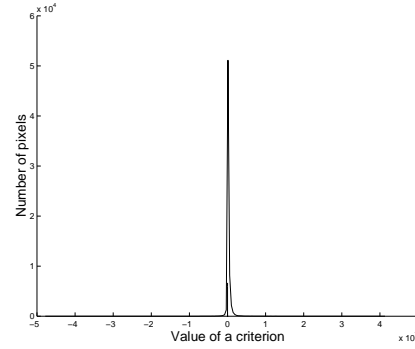


Fig. 5.11: Histogram of $\nabla u_\sigma \nabla v_\rho$

The aim of the next experiment is to demonstrate the advantage of the criterion $\nabla u_\sigma \nabla v_\rho$ over $|\nabla u_\sigma|^2$ in the case when boundary of objects in an image are blurred and not clearly defined. During the test, we consider the image presented in Fig. 5.12.

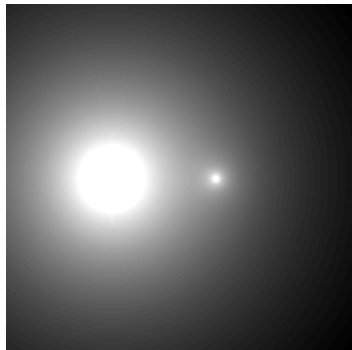


Fig. 5.12: Input image u

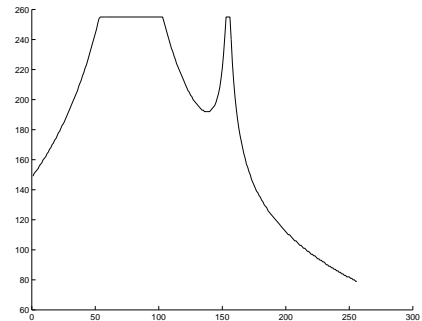


Fig. 5.13: Cross-section of u

In Fig. 5.14 we see the edge maps obtained by thresholding of criterion $|\nabla u_\sigma|^2$ for three selected values of the threshold μ . The same parameters have been used to obtain edge maps for criterion $\nabla u_\sigma \nabla v_\rho$. The results are presented in Fig. 5.15. As we can see in Fig. 5.14, it is difficult or even impossible to choose a value of the threshold μ , in way such that the boundaries of objects in the image u would be detected with required precision. In Fig. 5.14 c, one can see that the edge of the smaller circle is too thick and the edge of the bigger circle is not detected at all. Alternatively, the image presented in Fig. 5.15 c, carried out with the same threshold value $\mu = 15$ but for criterion $\nabla u_\sigma \nabla v_\rho$, gives more or less expected results.

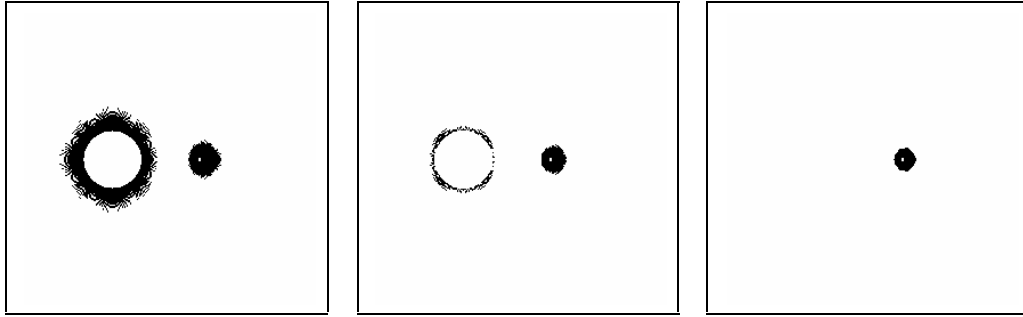


Fig. 5.14: Edge map of $|\nabla u_\sigma|^2$ with $\sigma = 0.1$ and the threshold a) $\mu = 5$, b) $\mu = 10$, c) $\mu = 15$

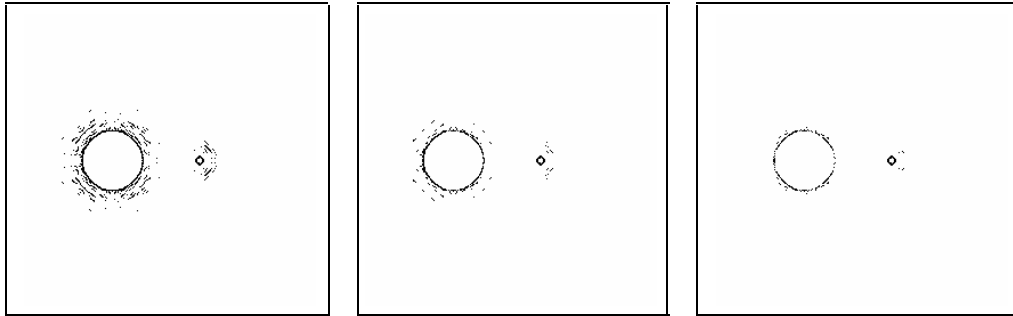
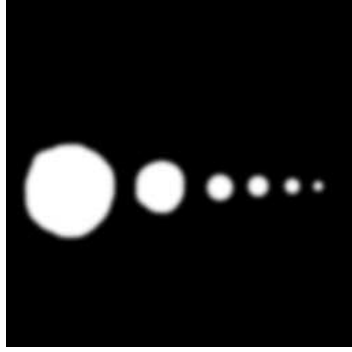
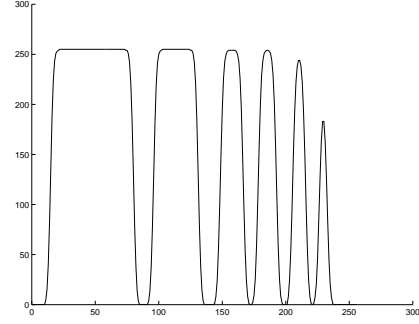
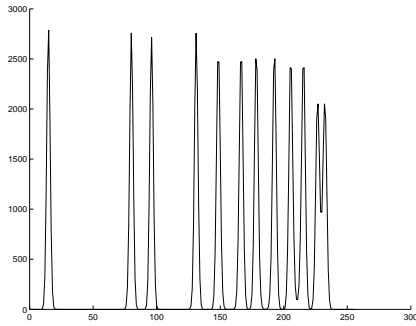
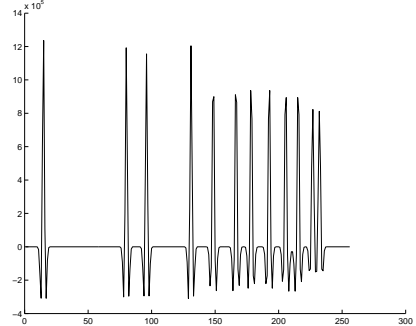


Fig. 5.15: Edge map of $\nabla u_\sigma \nabla v_\rho$ with $\sigma = \rho = 0.1$ and the threshold a) $\mu = 5$, b) $\mu = 10$, c) $\mu = 15$

In the third experiment, we would like to demonstrate results of edge detection with respect to choice of various threshold values. During the test, we calculate criteria $|\nabla u_\sigma|^2$ and $\nabla u_\sigma \nabla v_\rho$ with parameters $\sigma = \rho = 0.1$ for the image u presented in Fig. 5.16.

Fig. 5.16: Input image u Fig. 5.17: Cross-section of u

In Fig. 5.18 and 5.19 we present cross-sections of the both criteria.

Fig. 5.18: Cross-section of $|\nabla u_\sigma|^2$ Fig. 5.19: Cross-section of $\nabla u_\sigma \nabla v_\rho$

We observe that the values of criterion $|\nabla u_\sigma|^2$ in the position of potential edges are much greater than values of criterion $\nabla u_\sigma \nabla v_\rho$ at these positions. Furthermore, points for which the criterion is greater than 0 are more close to the center of edges. We should also notice the criterion $\nabla u_\sigma \nabla v_\rho$ is less sensitive to the choice of the threshold value. In Fig. 5.21, we can observe that for three different values of μ , the edge maps are nearly the same and the threshold from the interval $[0, 8 \cdot 10^5]$ gives relatively good and more or less similar results. Though, it should be noted that in the case of criterion $|\nabla u_\sigma|^2$, a reasonable value of a threshold can be chosen only from the interval $[10^3, 2 \cdot 10^3]$.

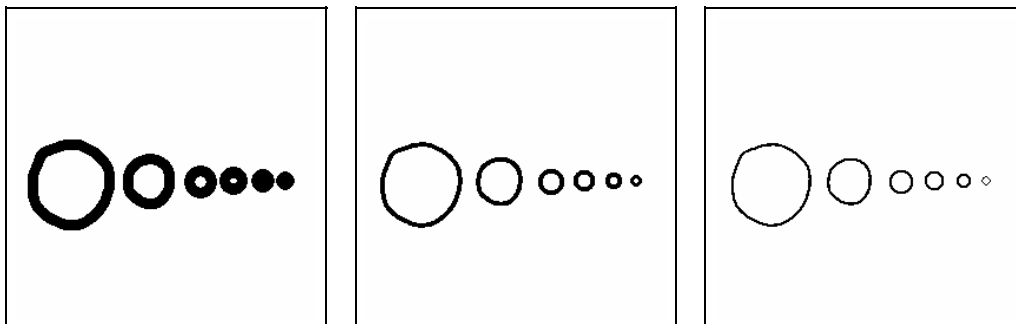


Fig. 5.20: Edge map of $|\nabla u_\sigma|^2$ with $\sigma = 0.1$ and the threshold a) $\mu = 10$, b) $\mu = 1000$, c) $\mu = 2000$

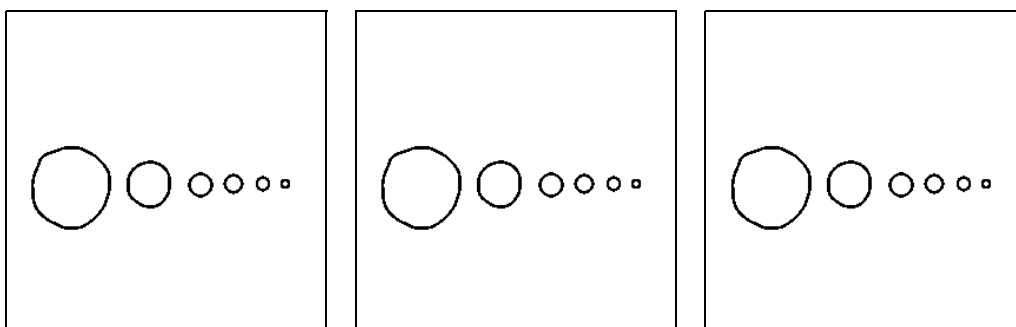


Fig. 5.21: Edge map of $\nabla u_\sigma \nabla v_\rho$ with $\sigma = \rho = 0.1$ and the threshold a) $\mu = 10$, b) $\mu = 1000$, c) $\mu = 2000$

5.2 Nonlinear isotropic diffusion filtering

In this section, we present results of isotropic diffusion filtering obtained by the application of the two semi-implicit finite volume schemes introduced in Section 4.2 of Chapter 4 for discretization of the Catté et al. model (2.8). In every discrete time step, each of these schemes result in a linear system of equations with a symmetric and strictly diagonally dominant matrix, which guarantees existence of a unique solution. The number of unknowns in such a system correspond to the number of cells in a grid. In the case of the adaptive filtering, in each discrete time step, a grid is adapted to the local image structure and a number of unknowns in the linear system is rapidly reduced. In the case of nonadaptive grid, the number of cells is the same in each iteration and is equal to the number of pixels in an image. In all examples, we work with the diffusivity d defined by

$$d(s) := \frac{1}{1 + s/\mu}$$

We consider two cases for each experiment: $s = |\nabla u_\sigma|^2$ and $s = \nabla u_\sigma \nabla v_\rho$. The regularizing convolution with the Gaussian kernel is implemented via the implicit finite volume scheme for linear diffusion.

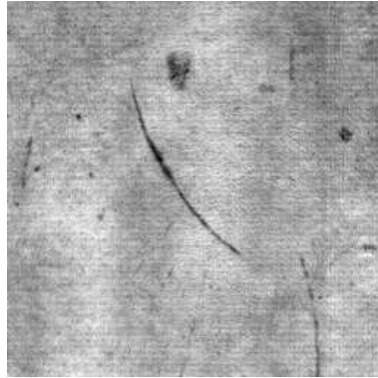


Fig. 5.22: Input image $f \in [0, 255]$, $\Omega = [0, 256]^2$

In Fig. 5.23, we present results of the nonadaptive isotropic diffusion filtering applied to the image presented in Fig. 5.22, using $|\nabla u_\sigma|^2$ with the threshold $\mu = 20$, and $\nabla u_\sigma \nabla v_\rho$ with the threshold $\mu = 50$ as an edge detector.

During the experiment, we have used the following parameters $\sigma = \rho = 1$, $\tau = 1$ and $iter = 20$. In both cases, the diffusion coefficient on the interface of two neighboring cells is approximated by arithmetic averaging.

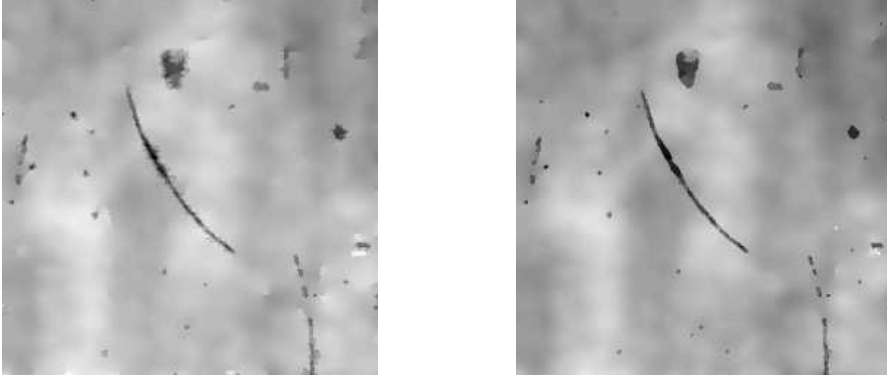


Fig. 5.23: Result of the isotropic diffusion filtering with the edge detector a) $|\nabla u_\sigma|^2$, b) $\nabla u_\sigma \nabla v_\rho$

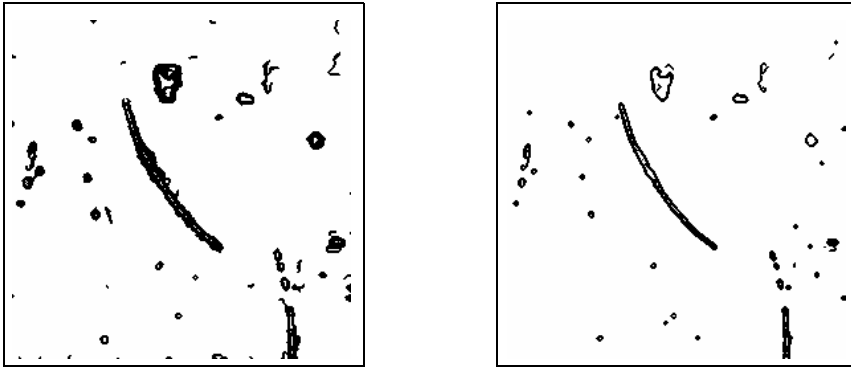


Fig. 5.24: Edge map obtained by thresholding of a) $|\nabla u_\sigma|^2$, b) $\nabla u_\sigma \nabla v_\rho$

In Fig. 5.24, we present edge maps obtained by thresholding of the criteria $|\nabla u_\sigma|^2$ and $\nabla u_\sigma \nabla v_\rho$ after the 20th iteration of the diffusion process. We observe that the position of edges in Fig. 5.24 b is more accurate and close to the center of the real edges. Owing to this, the boundaries obtained in the result of filtering are more sharp, as presented in Fig. 5.23 b. Moreover, there is less noise preserved at position of edges compare to the result presented in Fig. 5.23 a. Similar imperfection could be observed in Fig. 2.5, in Chapter 2.

In Fig. 5.25, we see the results of adaptive isotropic diffusion filtering using semi-implicit finite volume scheme with the diffusion coefficient discretized using arithmetic averaging. We have done two experiments. In the first, we used $|\nabla u_\sigma|^2$ as the edge detector with threshold $\mu = 30$ and in the second, $\nabla u_\sigma \nabla v_\rho$ with threshold $\mu = 100$. For both experiments, we used parameters $\sigma = \rho = 1$, $\tau = 1$ and $iter = 10$.

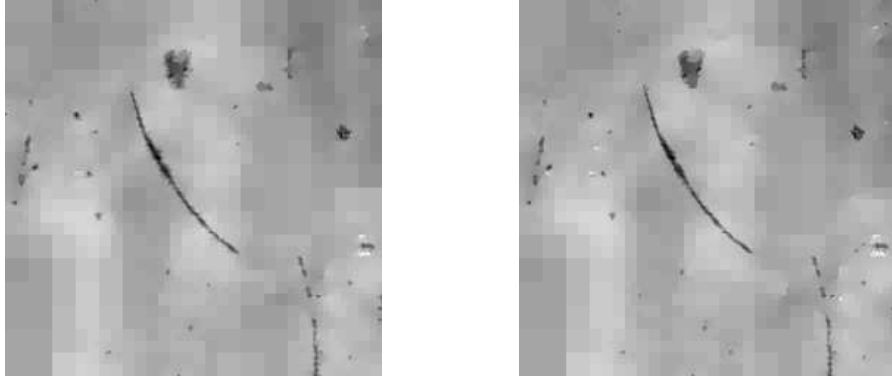


Fig. 5.25: Result of the adaptive isotropic diffusion filtering with arithmetic averaging of the diffusivity and the edge detector a) $|\nabla u_\sigma|^2$, b) $\nabla u_\sigma \nabla v_\rho$

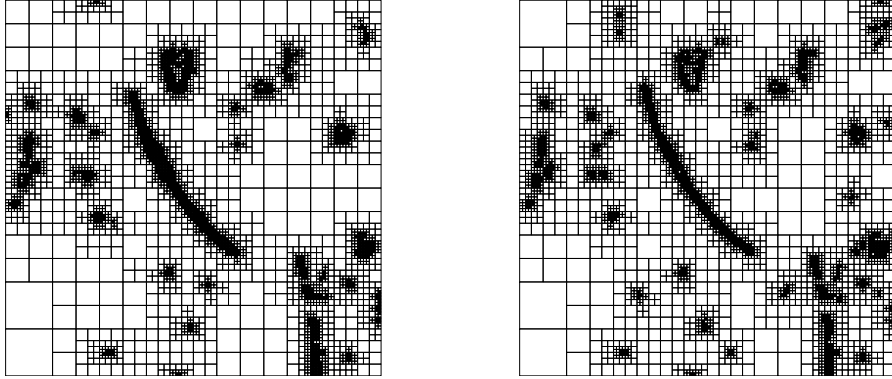


Fig. 5.26: Adaptive grid corresponding to the results presented in Fig. 5.25 with a) 6037 cells, b) 6094 cells

In Fig. 5.26, we present adaptive grids corresponding to the results presented in Fig. 5.25. Note that the input image used in filtering has size 256×256 . This means, that by the application of adaptive coarsening strategy, we were

able to reduce the number of grid cells and unknowns in the linear system, from 65536 to around 6000.

The results carried out with the same parameters as in the previous experiment, but for the case of the harmonic averaging of the diffusion coefficient on interface of two cells are presented in Fig. 5.27

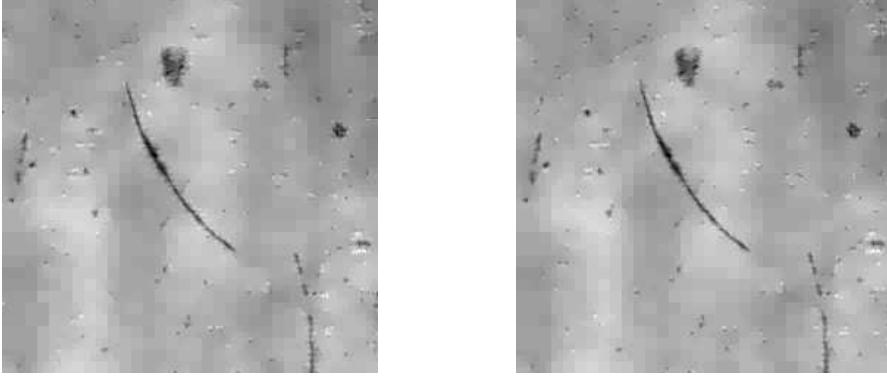


Fig. 5.27: Result of the adaptive isotropic diffusion filtering with harmonic averaging of the diffusivity and the edge detector a) $|\nabla u_\sigma|^2$, b) $\nabla u_\sigma \nabla v_\rho$

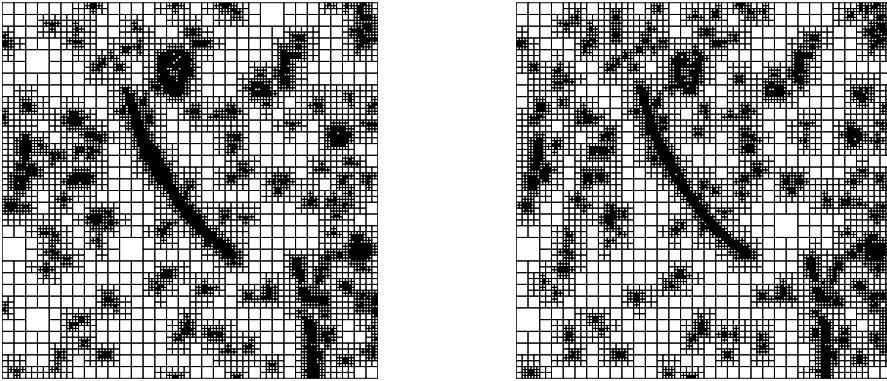


Fig. 5.28: Adaptive grid corresponding to the result presented in Fig. 5.27 with a) 10300 cells, b) 10156 cells

In Fig. 5.28, we present adaptive grids corresponding to the results presented in Fig. 5.27. In Tab. 5.1, we see the number of cells for each iteration of the adaptive isotropic diffusion filtering with respect to kind of averaging and used edge detector.

Iteration	0	1	2	3	4	5	6	7
B	65536	25987	14851	10186	8086	7276	6778	6433
A	65536	19288	12790	10024	8668	7951	7444	6925
C	65536	25987	17521	14620	13237	12214	11632	11284
D	65536	19288	14515	12961	12148	11614	11194	10780
Iteration	8	9	10					
B	6250	6124	6037					
A	6517	6298	6094					
C	10969	10714	10300					
D	10576	10396	10156					

Tab. 5.1: Number of cells for each iteration of the adaptive isotropic diffusion filtering. A: $|\nabla u_\sigma|^2$ and arithmetic averaging, B: $\nabla u_\sigma \nabla v_\rho$ and arithmetic averaging, C: $|\nabla u_\sigma|^2$ and harmonic averaging, D: $\nabla u_\sigma \nabla v_\rho$ and harmonic averaging

Chapter 6

Conclusions

In this thesis, the problem of nonlinear diffusion filtering of images with preserving and enhancing edges was theoretically and numerically investigated. We considered the topological asymptotic expansion of the functional J_ε , defined in (3.91), with respect to change the diffusivity function α_ε , defined in (3.5). We have proven that the asymptotic expansion

$$J_\varepsilon(u_\varepsilon) - J_0(u_0) = \varepsilon^2 g(x) + o(\varepsilon^2)$$

holds, and have derived the explicit form of its dominant term g for two different shapes of the inhomogeneity $B_\varepsilon = \varepsilon B$. In the case B is a disk of radius a , the explicit formula of the topological gradient g is given by

$$g(x) = (\alpha - 1) \frac{\alpha}{1 + \alpha} \pi a^2 |\nabla u_0|^2$$

In the case when B is an arbitrary ellipse with semi-axes a and b , where $a < b$, the topological gradient g is the matrix defined in (3.86). It has been shown that g is minimal if and only if the eigenvector corresponding to the greatest eigenvalue of this matrix is normal to the edge. Then, the value of g can be calculated using the following formula

$$g(x) = (\alpha - 1) \frac{\alpha(a + b)}{a + \alpha b} \pi ab |\nabla u_0|^2$$

Recalling the solution of equation $u_0 - f = \Delta u_0$ can be regarded as an implicit time discretization of the diffusion process with a single time step of size 1, we can generalize the definition of g and instead of $|\nabla u_0|^2$ take $|\nabla u_\sigma|^2$, where $u_\sigma := G_\sigma * u$.

Moreover, we have proposed new (to knowledge of the author) formula for a criterion that can be used to edge detection. This formula is given by

$$\nabla u_\sigma \nabla v_\rho$$

where

$$v_\rho := G_\rho * (-\langle \nabla u_\sigma, \nabla^2 u_\sigma \nabla u_\sigma \rangle)$$

We remark that if we would replace $|\nabla u_\sigma|^2$ in (2.8) by $\nabla u_\sigma \nabla v_\rho$, it would not change the well-posedness of this problem (see detailed proof of Theorem 2.1 in Catté et al. [18]). The advantages of this criterion in application to edge detection are confirmed based on numerical experiments in Chapter 5.

Furthermore, we have proposed the finite volume discretization for the Catté et al. (2.8) and the Weickert (2.9) models. The proposed discretization is based on the integro-interpolation method introduced by Samarskii in [56]. The numerical schemes have been derived for the case of uniform and nonuniform cell-centered grids of the computational domain $\Omega \subset \mathbb{R}^2$. In order to generate a nonuniform grid we applied the adaptive coarsening technique. In Chapter 5, numerical results of application derived schemes for a nonlinear isotopic diffusion filtering are presented.

Appendix A

Definitions and fundamental theorems

Theorem A.1 (Gauss-Green theorem) *Let $u \in C^1(\Omega)$. Then*

$$\int_{\Omega} \frac{\partial u}{\partial x_i} dx = \int_{\partial\Omega} u n^i ds \quad (i = 1, \dots, N),$$

where n^i is the outward unit normal of $\partial\Omega$.

Theorem A.2 (Integration by parts formula) *Let $u, v \in C^1(\Omega)$. Then*

$$\int_{\Omega} \frac{\partial u}{\partial x_i} v dx = - \int_{\Omega} u \frac{\partial v}{\partial x_i} dx + \int_{\partial\Omega} u v n^i ds \quad (i = 1, \dots, N),$$

where n^i is the outward unit normal of $\partial\Omega$.

Theorem A.3 (Green's formulas) *Let $u, v \in C^2(\Omega)$. Then*

(i)

$$\int_{\Omega} \Delta u dx = \int_{\partial\Omega} \frac{\partial u}{\partial n} ds$$

(ii)

$$\int_{\Omega} \nabla u \cdot \nabla v dx = - \int_{\Omega} u \Delta v dx + \int_{\partial\Omega} \frac{\partial v}{\partial n} u ds$$

(iii)

$$\int_{\Omega} u \Delta v - v \Delta u \, dx = \int_{\partial\Omega} u \frac{\partial v}{\partial n} \, ds - \int_{\partial\Omega} v \frac{\partial u}{\partial n} \, ds$$

Theorem A.4 (Hölder inequality) *Assume $1 \leq p, q \leq \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then, if $u \in L^p(\Omega)$ and $v \in L^q(\Omega)$, we have*

$$\int_{\Omega} |uv| \, dx \leq \|u\|_{L^p(\Omega)} \|v\|_{L^q(\Omega)}$$

Theorem A.5 (Minkowski inequality) *Assume $1 \leq p \leq \infty$ and $u, v \in L^p(\Omega)$. Then*

$$\|u + v\|_{L^p(\Omega)} \leq \|u\|_{L^p(\Omega)} + \|v\|_{L^p(\Omega)}$$

Theorem A.6 (Poincaré inequality) *Let $\Omega \subset \mathbb{R}^N$ be a bounded open domain and $u \in W_0^{1,p}(\Omega)$ with $1 \leq p \leq N$. Then*

$$\|u\|_{L^q} \leq C \|\nabla u\|_{L^p} \quad \text{with } q \in \left[1, \frac{Np}{N-p}\right]$$

for some constant C depending only on p, N, q and Ω .

Definition A.1 (Continuous embedding) *Let V, W be Banach spaces. We say V is continuously embedded in W and write*

$$V \hookrightarrow W$$

if $V \subset W$ and there exist a constant C such that for all $u \in V$

$$\|u\|_W \leq C \|u\|_V$$

Theorem A.7 (Sobolev embedding theorem) *Let $\Omega \subset \mathbb{R}^N$ be a bounded open domain with C^1 boundary and let $1 \leq p \leq \infty$. We have*

$$(i) \quad \text{If } 1 \leq p \leq N \quad \text{then} \quad W^{1,p}(\Omega) \hookrightarrow L^q(\Omega), \quad \frac{1}{q} \geq \frac{1}{p} + \frac{1}{N}$$

- (ii) If $p = N$ then $W^{1,p}(\Omega) \hookrightarrow L^q(\Omega)$, $1 \leq q < \infty$
 (iii) If $p > N$ then $W^{1,p}(\Omega) \hookrightarrow C(\bar{\Omega})$

Definition A.2 Let V be a Banach space with norm $\|\cdot\|_V$. A bilinear form

$$a : V \times V \rightarrow \mathbb{R}$$

is called continuous if there exists $\alpha > 0$ with

$$|a(u, v)| \leq \alpha \|u\|_V \|v\|_V \quad \forall u, v \in V \quad (\text{A.1})$$

a is called coercive if there exists a constant $\beta > 0$ with

$$a(u, u) \geq \beta \|u\|_V^2 \quad \forall u \in V \quad (\text{A.2})$$

Theorem A.8 (Lax-Milgram theorem) Let V be a Hilbert space and let $a : V \times V \rightarrow \mathbb{R}$ be a continuous, coercive bilinear form, i.e. (A.1) and (A.2) hold. Then for any bounded linear functional l on V the variational equation

$$a(u, v) = l(v)$$

has unique solution $u \in V$ for all $v \in V$

Theorem A.9 (Riesz representation theorem) Let V be a Hilbert space and let $l \in V'$. Then there exist a unique $v \in V$ such that $l(v) = \langle u, v \rangle_V$ for all $v \in V$.

Definition A.3 (Well-posedness) When a minimization problem or a PDE admits a unique solution that depends continuously on the data, we say that the minimization problem or the PDE is well-posed in the sense of Hadamard. If existence, uniqueness, or continuity fails, we say that the minimization problem or the PDE is ill-posed.

List of notations

\mathbb{R}	Real numbers
\mathbb{R}_+	Positive real numbers, $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x > 0\}$
\mathbb{N}	Natural numbers, $\mathbb{N} = \{1, 2, 3, \dots\}$
\mathbb{R}^N	N -dimensional Euclidean space
$\mathbb{S}^{N \times N}$	Space of N -dimensional real symmetric matrices
Ω	Domain in \mathbb{R}^N
$\bar{\Omega}$	Closure of the domain Ω
$\partial\Omega$	Boundary of the domain Ω
∇u	Gradient of u in the classical sense $\nabla u = \left(\frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_N} \right)$
$\nabla \cdot u$	Divergence operator $\nabla \cdot u = \sum_{i=1}^N \frac{\partial u}{\partial x_i}$
$\nabla^2 u$	Hessian matrix of u in the classical sense $(\nabla^2 u)_{i,j} = \frac{\partial^2 u}{\partial x_i \partial x_j}$
Δu	Laplacian operator $\Delta u = \sum_{i=1}^N \frac{\partial^2 u}{\partial x_i^2}$
a^T	Vector transposed to a vector $a = (a_1, \dots, a_N) \in \mathbb{R}^N$
A^T	Matrix transposed to a matrix A
a^\perp	Vector perpendicular to a vector a
I_0	Modified Bessel function of the first kind
K_0	Modified Bessel function of the second kind
Φ	Fundamental solution of the Laplace equation $\Delta u = 0$ defined by $\Phi(x, y) := -\frac{1}{2\pi} \ln(x - y)$ for $x, y \in \mathbb{R}^2$ and $x \neq y$
Γ	Fundamental solution of the equation $u - \Delta u = 0$ defined by $\Gamma(x, y) := \frac{1}{2\pi} K_0(x - y)$ for $x, y \in \mathbb{R}^2$ and $x \neq y$
G_σ	Gaussian kernel defined by $G_\sigma(x, y) := \frac{1}{2\pi\sigma} \exp(-\frac{ x-y }{2\sigma^2})$ for $x, y \in \mathbb{R}^2$
$f * g$	Convolution of f and g
$\ln x$	Natural logarithm of x
$ x $	Absolute value of x
$\ u\ _V$	Norm of u in space V

$C_0^p(\Omega)$	Space of real valued functions, p continuously differentiable with compact support
$C_0^\infty(\Omega)$	Space of real valued functions, infinitely continuously differentiable with compact support
$L^p(\Omega)$	Space of Lebesgue measurable functions u such that

$$\|u\|_{L^p(\Omega)} := \left(\int_{\Omega} |u|^p dx \right)^{1/p} < \infty \quad \text{for } 1 \leq p \leq \infty$$

$L^\infty(\Omega)$	Space of Lebesgue measurable functions u such that
--------------------	--

$$\|u\|_{L^\infty(\Omega)} := \text{ess sup}_{\Omega} |u| < \infty$$

$L_{loc}^p(\Omega)$	$\{u : \Omega \rightarrow \mathbb{R} \mid u \in L^p(U) \text{ for each } U \subset\subset \Omega\}$
---------------------	---

$W^{k,p}(\Omega)$	With $1 \leq p < \infty$. Sobolev space of functions $u \in L^p(\Omega)$ such that for each multiindex α with $ \alpha \leq k$, derivative D^α exist in the weak sense and belongs to $L^p(\Omega)$. In $W^{k,p}(\Omega)$ we define the norm
-------------------	--

$$\|u\|_{W^{k,p}(\Omega)} := \left(\sum_{|\alpha| \leq k} \int_{\Omega} |D^\alpha u|^p dx \right)^{1/p}$$

In particular, we write $H^1(\Omega) = W^{1,2}(\Omega)$. Thus

$$\|u\|_{H^1(\Omega)} := \left(\int_{\Omega} u^2 dx + \int_{\Omega} Du \cdot Du dx \right)^{1/2}$$

$C([0, T]; L^2(\Omega))$	Space of continuous functions $u : [0, T] \rightarrow L^2(\Omega)$ with the norm
--------------------------	--

$$\|u\|_{C([0, T]; L^2(\Omega))} := \max_{[0, T]} \|u(t)\|_{L^2(\Omega)}$$

$L^2([0, T]; H^1(\Omega))$	Space of functions u measurable on $[0, T]$ for the Lebesgue measure dt with the range in $H^1(\Omega)$, such that
----------------------------	---

$$\|u\|_{L^2([0, T]; H^1(\Omega))} := \left(\int_0^T \|u(t)\|_{H^1(\Omega)}^2 dt \right)^{1/2} < \infty$$

Ω_m	Uniform grid composed of square cells $P \in \Omega_m$ with a side length $h_P = 2^{m-1}h$
Ω_{m+l}	Nonuniform grid composed of square cells $P \in \Omega_{m+l}$ with a side length $h_P \in \{2^{m-1}h, 2^{m+1}h, \dots, 2^l h\}$ for $m, l \in \mathbb{N}$ and $m < l$
$\mathcal{N}(P)$	Set of neighbours of the cell P
h_P	Side length of the cell P
$ P \cap Q $	Length of interface between cells P and Q , $ P \cap Q = \min(h_P, h_Q)$
(x_1^P, x_2^P)	Center of the cell P
μ	Threshold
σ, ρ	Standard deviations
τ	Time discretization step
$iter$	Number of iterations

Bibliography

- [1] G. Allaire, F. D. Gournay, F. Jouve, and A. M. Toader. Structural optimization using topological and shape sensitivity via a level set method. *Control and Cybernetics*, 34:59–80, 2005.
- [2] L. Alvarez, F. Guichard, P. L. Lions, and J. M. Morel. Axioms and fundamental equations of image processing. *Archive for rational mechanics and analysis*, 123(3):199–257, 1993.
- [3] L. Alvarez, P. L. Lions, and J. M. Morel. Image selective smoothing and edge detection by nonlinear diffusion. II. *SIAM journal on numerical analysis*, 29(3):845–866, 1992.
- [4] H. Amann. Time-delayed Perona-Malik equations. *Acta Mathematica Universitatis Comenianae*, LXXVI:15–38, 2007.
- [5] H. Ammari and H. Kang. Properties of the generalized polarization tensors. *Multiscale Modeling and Simulation*, 1(2):335–348, 2003.
- [6] S. Amstutz. *Aspects théoriques et numériques en optimisation de forme topologique*. PhD thesis, Laboratoire: Mathématiques pour l’Industrie et la Physique, Toulouse, 2004.
- [7] S. Amstutz. Sensitivity analysis with respect to a local perturbation of the material property. *RICAM report*, 24, 2005.
- [8] S. Amstutz and H. Andrä. A new algorithm for topology optimization using a level-set method. *Journal of Computational Physics*, 216:573–588, 2006.
- [9] S. Amstutz, B. Samet, and N. Dominguez. Sensitivity analysis with respect to the insertion of small inhomogeneities. *Proceedings of ECCO-MAS*, 2004.

-
- [10] G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations (second edition)*, volume 147 of *Applied Mathematical Sciences*. Springer-Verlag, 2006.
 - [11] D. Auroux, M. Masmoudi, and L. Belaid. Image restoration and classification by topological asymptotic expansion. *Variational Formulations in Mechanics: Theory and Applications*, 2006.
 - [12] E. Bänsch and K. Mikula. A coarsening finite element strategy in image selective smoothing. *Computing and Visualization in Science*, 1:53–61, 1997.
 - [13] E. Bänsch and K. Mikula. Adaptivity in 3d image processing. *Computing and Visualization in Science*, 4:21–30, 2001.
 - [14] C. A. Bouman and K. D. Sauer. A generalized gaussian image model for edge-preserving MAP estimation. *IEEE Transactions on Image Processing*, 2(3):296–310, 1993.
 - [15] M. Brühl, M. Hanke, and M. S. Vogelius. A direct impedance tomography algorithm for locating small inhomogeneities. *Numerische Mathematik*, 93(4):635–654, 2003.
 - [16] M. Burger, B. Hackl, and W. Ring. Incorporating topological derivatives into level set methods. *Journal of Computational Physics*, 194:344–368, 2004.
 - [17] Y. Capdeboscq and M. S. Vogelius. A review of some recent work on impedance imaging for inhomogeneities of low volume fraction. *Contemporary Mathematics*, 362:69–88, 2004.
 - [18] F. Catte, P. L. Lions, J. M. Morel, and T. Coll. Image selective smoothing and edge detection by nonlinear diffusion. *Journal of Numerical Analysis*, 29(1):182–193, 1992.
 - [19] J. C  a. Conception optimale ou identification de forme, calcul rapide de la d  riv  e directionnelle de la fonction cout. *M.A.A.N.*, 20(3):371–401, 1986.
 - [20] J. C  a, S. Garreau, P. Guillaume, and M. Masmoudi. The shape and topological optimizations connection. *Computer methods in applied mechanics and engineering*, 188(4):713–726, 2000.

-
- [21] D. J. Cedio-Fengya, S. Moskow, and M. S. Vogelius. Identification of conductivity imperfections of small diameter by boundary measurements. continuous dependence and computational reconstruction. *Inverse Problems*, 14:553–595.
 - [22] T. F. Chan and J. J. Shen. *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*. Society for Industrial and Applied Mathematics, 2005.
 - [23] D. Colton and A. Kirsch. A simple method for solving inverse scattering problems in the resonanse region. *Inverse Problems*, 12(4):383–393, 1996.
 - [24] D. Colton and R. Kress. *Integral Equation Methods in Scattering Theory*. Wiley, New York, 1983.
 - [25] D. Colton and R. Kress. *Inverse Acoustic and Electromagnetic Scattering Theory*. Springer, 1998.
 - [26] G. Demoment. Image reconstruction and restoration: Overview of common estimation structures and problems. *IEEE Trans. Acoustics, Speech and Signal Processing*, 37(12):2024–2036, 1989.
 - [27] D. L. Donoho. Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition. *Applied and Computational Harmonic Analysis*, 2:101–126, 1995.
 - [28] D. L. Donoho and I. M. Johnstone. Adapting to unknown smoothness by wavelet shrinkage. *Journal of the American Statistical Association*, 90:1200–1224, 1995.
 - [29] L. C. Evans. *Partial Differential Equations*. AMS, 1998.
 - [30] R. Ewing, R. Lazarov, and P. Vassilevski. Local refinement techniques for elliptic problems on cell-centered grids. I: Error analysis. *Math. Comp.*, 56:437–461, 1991.
 - [31] R. Ewing, R. Lazarov, and P. Vassilevski. Local refinement techniques for elliptic problems on cell-centered grids. II: Optimal order two-grid iterative methods. *Numer. Linear Algebra Appl.*, 1:337–368, 1994.
 - [32] G. R. Feijóo. A new method in inverse scattering based on the topological derivative. *Inverse Problems*, 20:1819–1840, 2004.

-
- [33] R. A. Feijóo, A. Novotny, C. Padra, and E. Taroco. The topological derivative for the Poisson problem. *Mathematical Models and Methods in Applied Sciences*, 13(12):1825–1844, 2003.
 - [34] V. Friederichs. *Local Smoothing Methods in Image Processing*. PhD thesis, University of Kaiserslautern, 1999.
 - [35] S. Garreau, P. Guillaume, and M. Masmoudi. The topological asymptotic for PDE systems: The elasticity case. *SIAM Journal on Control and Optimization*, 39(6):1756–1778, 2000.
 - [36] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):101–126, 1995.
 - [37] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Addison-Wesley, 3rd edition, 1992.
 - [38] P. Ion, W. Hackbusch, and R. Fadiman. *Elliptic Differential Equations*. Springer, 1992.
 - [39] J. Jost and X. Li-Jost. *Calculus of Variations*. Cambridge University Press, 1999.
 - [40] H. Kang and K. Kim. Anisotropic polarization tensors for ellipses and ellipsoids. *Jour. Comp. Math.*, to appear.
 - [41] J. J. Koenderink. The structure of images. *Biological Cybernetics*, 1984.
 - [42] R. Kress. *Linear Integral Equations*. Springer, 1989.
 - [43] Z. Krivá and K. Mikula. An adaptive finite volume method in image processing of color images. *Proceedings of Algoritmy, Conference on Scientific Computing*, 2000.
 - [44] Z. Krivá and K. Mikula. An adaptive finite volume scheme for solving nonlinear diffusion equations in image processing. *Journal of Visual Communication and Image Representation*, 2000.
 - [45] I. Larrabide, R. A. Feijóo, A. A. Novotny, E. Taroco, and M. Masmoudi. An image segmentation method based on a discrete version of the topological derivative. *6th World Congress on Structural and Multidisciplinary Optimization*, 2005.

-
- [46] D. Marr and E. Hildreth. Theory of edge detection. *Proceedings of the Royal Society of London*, pages 187–217, 1980.
 - [47] K. Mikula and N. Ramarosy. Semi-implicit finite volume scheme for solving nonlinear diffusion equations in image processing. *Numerical Mathematic*, 89:561–590, 2001.
 - [48] L. Moisan, F. Guichard, and J. Morel. A review of PDE models in image processing and image analysis.
 - [49] J. M. Morel and S. Solimini. *Variational Methods in Image Segmentation*. Birkhäuser, Boston, 1995.
 - [50] P. Mrázek, J. Weickert, and A. Bruhn. On robust estimation and smoothing with spatial and tonal kernels. pages 335–352, 2006.
 - [51] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and variational problems. *Comm. on Pure and Applied Math.*, XLII(5):577–685, 1988.
 - [52] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *Proc. IEEE Computer Soc. Workshop on Computer Vision*, pages 16–27, 1987.
 - [53] A. D. Polyanin. *Handbook of Linear Partial Differential Equations*. Chapman & Hall/CRC, 2002.
 - [54] T. Preusser and M. Rumpf. An adaptive finite element method for large scale image processing. *Proceedings of Scale Space*, pages 223–234, 1999.
 - [55] G. F. Roach. *Green's Functions*. Cambridge University Press, 1982.
 - [56] A. A. Samarskii. *The Theory of Difference Schemes*. CRC, 2001.
 - [57] A. Schumacher. *Topologieoptimierung von Bauteilstrukturen unter Verwendung von Topchpositionierungskriterien*. PhD thesis, Universität-Gesamthochschule-Siegen, 1995.
 - [58] J. Sokołowski and A. Żochowski. On topological derivative in shape optimization. *SIAM Journal on Control and Optimization*, 37, 1999.
 - [59] J. Sokołowski and A. Żochowski. Topological derivatives for elliptic problems. *Inverse Problems*, 15:123–134, 1999.
 - [60] J. Spanier. *An Atlas of Functions*. Taylor & Francis, 1987.

- [61] M. S. Vogelius and D. Volkov. Asymptotic formulas for perturbations in the electromagnetic fields due to the presence of inhomogeneities of small diameter. *Mathematical Modeling and Numerical Analysis*, 34(4):723–748, 2000.
- [62] J. Weickert. A review of nonlinear diffusion filtering. *Lecture Notes in Computer Science*, 1252:3–28, 1997.
- [63] J. Weickert. *Anisotropic Diffusion in Image Processing*. Teubner Verlag, Stuttgart, 1998.
- [64] J. Weickert, B. Romeny, and M. Viergever. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Transactions on Image Processing*, 7(3):398–410, 1998.

CURRICULUM VITAE

- 3 July 1978 born in Głogów, Poland
- 1993 – 1997 I Lyceum in Głogów
- Okt. 1997 studies at Wrocław University of Technology
– Sep. 1998 Faculty of Fundamental Problems of Technology
Department of Mathematics
- Okt. 1998 studies at University of Zielona Góra
– Sep. 2003 Faculty of Science, Department of Mathematics
- Apr. 2003 ERASMUS program at University in Kaiserslautern
– Aug. 2003 Department of Industrial Mathematics
- Oct. 2003 PhD studies at Technical University in Kaiserslautern
– Apr. 2007 Department of Industrial Mathematics